

Hazard proximities of childhood cancers in Great Britain from 1953-80

E G Knox, E A Gilman

Abstract

Study objectives - Firstly, to examine relationships between the birth and death addresses of children dying from leukaemia and cancer in Great Britain, and the sites of potential environmental hazards; and secondly to measure relative case densities close to, and at increasing distances from, different hazard types.

Design - Home address postcodes (PCs) and their map coordinates were identified at birth and at death in children who died from leukaemia or cancer. Potentially hazardous industrial addresses and PCs were listed from business and other directories, and map coordinates obtained from the *Central Postcode Directory* or else located directly on Ordnance Survey (OS) maps. Railway lines and motorways were digitised from OS maps. Numbers of deaths (and births) at successive radial distances from these hazards were counted and compared with expected numbers. The latter were based on a count of all PCs at similar distances. Relative case density ratios at successive distances from the hazards were obtained from observed and expected numbers, aggregated over similar sites. This was repeated for different hazard types and results were tested for evidence of systematic centrifugal case density gradients.

Participants and setting - All 22 458 children dying from leukaemia or cancer aged 0-15 years, in England, Wales, and Scotland, between 1953 and 1980.

Main results - Relative excesses of leukaemias and of solid cancers were found near the following: (1) oil refineries, major oil storage installations, railside oil distribution terminals and factories making bitumen products; (2) motor car factories, coach builders, and car body repairers; (3) major users of petroleum products including manufacturers of solvents, paint sprayers, fibreglass fabricators, paint and varnish makers, plastics and detergent manufacturers, and galvanisers; (4) users of kilns and furnaces including steelworks, power stations, galvanizers, cement makers, brickworks, crematoria and aluminium, zinc, and iron/steel foundries; (5) airfields, railways, motorways and harbours. The findings for leukaemias and for solid cancers were indistinguishable. The hazard proximities of birth addresses were stronger than for death addresses. For children who had moved house between

birth and death, the proximity effect was limited to the birth addresses.

Conclusions - Childhood cancers are geographically associated with two main types of industrial atmospheric effluent namely: (1) petroleum derived volatiles and (2) kiln and furnace smoke and gases, and effluents from internal combustion engines.

(*J Epidemiol Community Health* 1997;51:151-159)

Our previous studies showed that childhood leukaemias and cancers occurred in small geographical clusters.^{1,2} Among 9411 childhood registrations of leukaemia and lymphoma in Great Britain between 1966 and 1983, for whom we knew the residential postcode (PC) coordinates, there were 264 case pairs/triplets who were separated by less than 150m; and in England and Wales there were 520 registration pairs who shared a common census enumeration district (ED). Among 22 458 childhood leukaemia and cancer deaths in Great Britain between 1953 and 1980, 503 pairs shared a common PC. All these numbers were about twice the expected values. The excesses were statistically significant, and they were not explained by national variations of ED and PC sizes (numbers of households and of postal delivery points). Studies of house numbering revealed a further excess of pairs which bridged the boundaries of adjacent PCs.

The close address clusters were about 300 m in diameter, but there was also a surrounding zone of moderately enhanced risk, tapering towards normal over a radius of 2-3 km. PC clustering was stronger among birth addresses than among death addresses. It was not limited to the leukaemias and it involved all childhood cancers. Some of the close pairs were concordantly affected siblings living at the same address, but their exclusion did not alter the overall conclusion. The scale of the geographical clusters was commensurate with that of previously demonstrated space-time clusters among the 1966-83 leukaemia registrations, and also among the births and onsets of the fatal 1953-80 leukaemias.^{3,4,5}

If the geographic clustering is genuine and not a demographic artefact, as is now clear, then it must reflect the existence of localised environmental hazards. Many of them must have persisted over periods of years, although the space-time interactions show that some could have been transient or intermittent. They act most powerfully within a few hundred metres, but some of them reach out to a few kilometres.

Mill Cottage,
Great Comberton,
Worcestershire WR10
3DU
E G Knox

Department of Public
Health and
Epidemiology,
Medical School,
University of
Birmingham,
Birmingham B15 2TT
E A Gilman

Correspondence to:
Professor E G Knox.

Accepted for publication
September 1996

A previous study showed that the map locations of the 9411 registrations were closer to a variety of potentially hazardous sites than were a randomly chosen set of PCs.⁶ The 264 sub-150m pairs among them showed an even stronger relationship. The apparent hazards included oil refineries, oil storage and distribution depots, railway lines, and other industrial sites. Effective ranges extended as far as 5 km from the sources. This suggested a hazard related to large scale usages of fossil fuels, especially petroleum, operating through leakage or evaporation or combustion – perhaps all three. The matching secular trends of childhood leukaemia and of petroleum usage over many decades accords with this hypothesis.

However, there were several problems of interpretation. First, the proximity investigations were undertaken initially to test the reality of the then uncertain clustering, and any causal speculations involved a second use of the same data. This introduced an unsatisfactory circularity. Second, no child population denominators were available for geographical analyses conducted on this small scale. The use of randomly selected PCs as indicators of the geographical distribution of those at risk entailed several demographic uncertainties. The most important was the possibility that numbers of children in PCs close to industrial sites might show a systematic positive bias. This caveat was also elaborated by commentators.⁷ Third, there was a question of environmental confounding; hazards of one kind might be related geographically to hazards of another kind, with a consequent risk of false attribution. This applied especially to railways. Finally, the close (300 m) spacing of the core clusters did not match the wider (c 5 km) radii of the main hazard proximity effects.

These problems demanded additional study. Like the previous one, this reinvestigation lacks a true measure of child exposure on the requisite fine scale. No such data are generally available. The fundamental problems of interpretation are therefore the same. However, it is based on a fresh data set – larger and more comprehensive, and including addresses and dates of birth as well as of onset and death. It records solid cancers as well as leukaemias. But most crucially it starts from the now well supported premise that persistent local hazards must indeed exist. This no longer needs to be demonstrated. The present task is simply to discover what they are.

Methods

The cancer data set used here is the second of the two noted above, and it has been described elsewhere.^{6,8,9} It records all 22 458 deaths from leukaemias and other cancers in children aged 0–15 years in England, Wales, and Scotland between 1953 and 1980. After excluding deaths in the Northern Isles of Scotland, and a few recording errors, we were left with a working file of 22 448 death addresses. The birth addresses of many were also available. Although these data are limited to fatal cases, the great majority of the children affected in these years

died within the time limits of the investigation; over 75% within five years of diagnosis.¹⁰

The existence and locations of putative environmental hazards were determined in several ways. Point hazards such as oil refineries, gasworks, airfields, and crematoria were identified through direct searches of Ordnance Survey (OS) maps and street atlases. Many factories were found in classified business directories.¹² Others such as motor car works, nuclear installations, benzene refineries, and TV transmitters were taken from specialised lists. The lists and directories supplied addresses and PCs, which were then translated to map references via the *Central Postcode Directory*. Where “business” PCs could not be traced in our “residential” directory we used the alphabetically nearest residential PC (last character) as the index position. Business addresses listed as “Company House” were avoided in favour of actual factories. Some locations were reconfirmed or amended through map inspections or site visits. Detected errors were typically between 100 and 300 m. Series of particular hazard types usually consisted of 30–50 locations taken from the index in alphabetic order. For a few uncertain results, extensions of these lists enabled us to reject a suspect finding.

Linear cartographic features – railways and motorways – were digitised from OS maps using a digitising tablet. This registered strings of point coordinates and allowed the labelling of their beginnings, ends, breaks, and branch points. Tablet coordinates were translated to map positions against the key punched positions of the map sheet corners. The railways were digitised in 11 separate regions, with London excluded because of its density and complexity. Motorways outside the recently completed M25 London Orbital (which was itself omitted) were digitised as a single set.

For the point hazards, numbers of cases and numbers of PCs were counted in successively greater circles around each source. Crude densities of cases per 1000 PCs within each annulus were calculated from the aggregated numbers, and many close proximity concentrations were found. However, we feared that higher case densities in some regions of the country may have coincided with geographical concentrations of certain industries. The calculations were therefore repeated using a form of indirect regional standardisation. This was based upon direct measurements of case densities in 56 urban and non-urban areas, as described later. Region specific expected numbers of cases, within successive annuli, were then calculated. The expected number was the product of the regional case density per PC and the local number of PCs within the annulus. The observed and expected results were summed to generate a standardised density ratio (SDR) for each successive circle around each hazard type. The method is set out in an appendix.

For linear map features (rail, road) we found and measured the shortest case to hazard distance. Each address was examined against every node in the net and perpendiculars were

Table 1 Residential postcode (PC) densities (per km²) at different distances from index points

Index	Distance (km)									
	(No)	0-0.3	-0.6	-1.0	-2.0	-3.0	-4.0	-5.0	-7.0	-10.0
Random PCs	(500)	94.05	70.42	60.96	47.83	39.81	35.99	33.21	30.28	26.67
Random cases	(700)	98.12	75.68	66.31	52.66	46.04	42.48	38.94	34.79	30.37
Refineries*	(21)	0.51	2.25	5.04	9.56	15.06	14.11	16.35	15.82	13.84
Airfields	(192)	2.43	3.74	6.62	8.20	9.72	10.13	10.19	10.32	9.37
Oil stores†	(149)	25.80	33.75	36.80	32.38	26.14	22.85	19.86	16.52	15.44
Chemical works‡	(83)	37.67	33.76	34.91	35.47	32.01	28.68	27.10	24.84	17.92
Crematoria	(64)	27.63	50.53	59.65	53.75	49.60	42.81	37.33	32.06	27.76
Car making	(193)	62.21	58.98	63.03	56.09	47.92	42.18	36.38	31.17	27.67
Safe sites?	(195)	105.21	88.48	77.99	67.06	57.10	51.35	47.72	41.65	42.90

* Including 6 benzene refineries.

† Combined major oil storage locations, and railside (mainly) oil distribution terminals.

‡ Combined "petrochemical" and other "industrial chemicals" works plus other factories noted to have storage tank facilities.

§ See text.

Total land area of England, Scotland and Wales is 229 870 km². The overall density of the 1 372 359 residential postcodes is 5.97/km².

dropped to inter-node segments. Cancer addresses were allocated to parallel bands on either side of the line rather than to annuli. A similar allocation of all PCs was impractical here, so numbers in different distance bands were estimated from the locations of 10000 randomly selected PCs. The counted numbers were adjusted according to the sampling fraction, and this was used in turn to calculate a regionally standardised expected number, as above.

In the absence of population based denominators for individual PCs, alternative small group analyses were considered. A reclassification of addresses by census ED could have supplied an alternative set of coordinates, and could have provided local estimates of child years at risk. Unfortunately, appropriate census ED data were not available for the greater part of the study period. This choice would also have discarded the superior geographical resolution of the PCs. There are approximately 11 PCs per ED in England and Wales, and other studies using census based child denominators found in practice that EDs had to be aggregated into even larger blocks. These studies also failed to show clustering at these relatively coarse levels.^{13,14} This approach was unlikely to meet the needs of a hazard proximity study pursuing very close range effects: and which was founded upon an initial demonstration of clustering within much smaller areas.

The matched case-control format of our main data source suggested another method for displaying comparable proximity patterns among non-cancer children. Unfortunately, control addresses had not been postcoded, matched controls were available only for about two thirds of cases, and they had been geographically matched according to place of residence at the time that the "case" had died. They were not thus really suitable for this task. Over matching in this respect was likely to obscure any true proximity effects. In pragmatic terms the extra work was beyond current resources and it seemed unlikely to attract them.

As with our previous proximity study, the crucial demographic question was whether apparent case concentrations might result entirely from systematic concentrations of PCs with large child populations around the types of hazard examined. Again, no appropriate data

were available within the distance ranges considered. Regional standardisation eliminates only one component of the problem and less resolved population groupings (eg EDs) could offer no advantages. This focal demographic artefact was entirely hypothetical, and unsupported by external evidence, but we felt that it should be considered seriously. Fortunately, we found ways of examining the issue within our own data.

Results

SHORT RANGE DEMOGRAPHIC EFFECTS

Crude case densities per 1000 PCs were measured in different parts of the country: 56 rectangular areas covering major cities and inner cities, and residual surrounding areas. They were greater in population-dense than in sparse areas. The mean overall density was 16.36 deaths per 1000 residential PCs, but individual values descended from Coveney (29.98), Scunthorpe (27.78), Doncaster (24.30), Glasgow (24.01), Liverpool (23.27), Hull (23.15), Birmingham (23.07), central London (22.41) ... to ... Southampton (14.32), Gloucester (13.49), Brighton (12.06), Exeter (11.16), and Bournemouth (9.34); and the residual non-urban zones of northern England, Wales, and southwest England (10.28, 9.60, 7.79).

Although correlations with high population density have been demonstrated elsewhere,¹⁴ these regional variations do not accurately follow known mortality patterns. They must stem in part from the different ways in which PC boundaries are delineated. PCs in population-dense industrial zones probably have more delivery points (DPs) per PC, and possibly more children per DP. Such excesses would create a false overall association between industrial hazards and cancer deaths per 1000 PCs. This necessitated the development of the stratification and standardisation procedures outlined above.

Unfortunately, broad area standardisations provide no protection against the falsifying effects of high DP PCs in the immediate vicinities of industrial premises. There is no available measure of industrially related demographic distributions on this scale. However, we tested the proposition indirectly by calculating densities of PCs per km² around each of the hazard types examined. The results are sum-

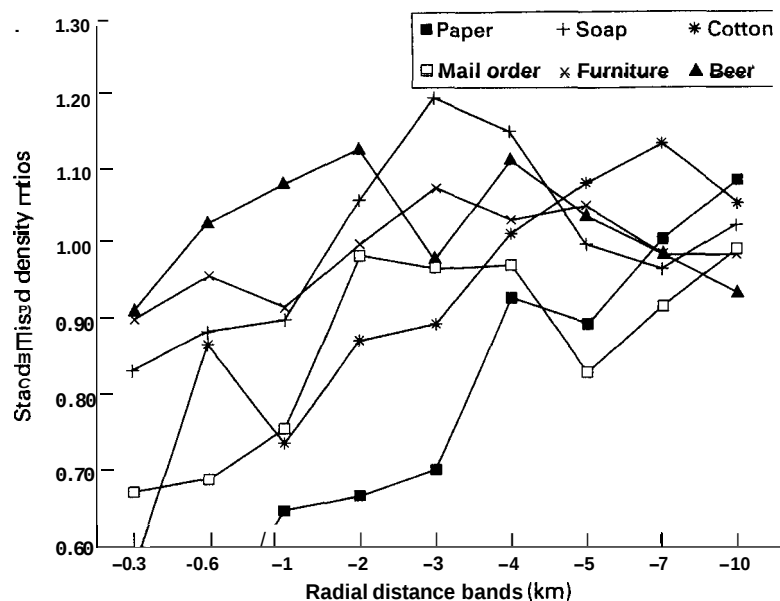


Figure 1 Industrial/residential competition.

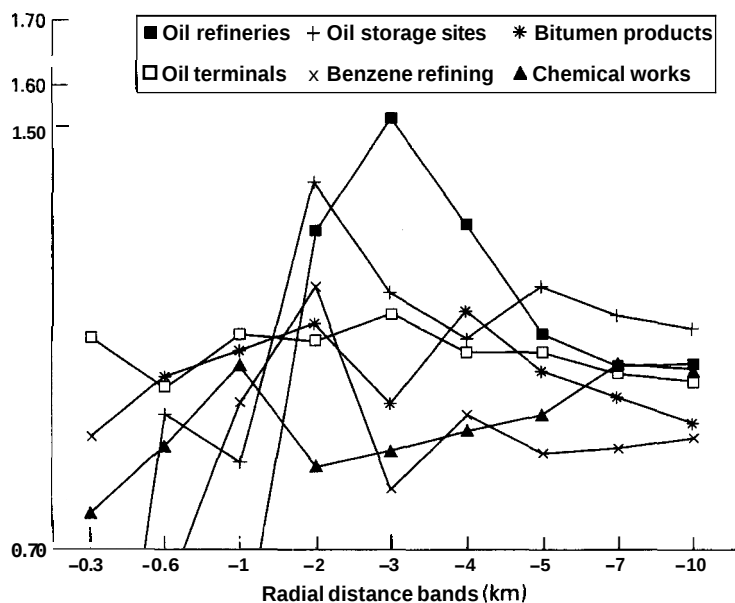


Figure 2 Oil refining and storage.

marised in table 1. PC densities are compared around:

- o A random sample of PCs;
- o A random sample of cancer deaths; and
- o Potential hazards of several different types.

Local PC densities around random PCs and around cancer cases were much greater than the overall national value (5.97 PCs/km²) based upon total land area. This is because the sample selection process is biased in favour of densely inhabited regions. At increasing distances, the high densities tapered towards the national value as this selection effect became diluted. The much lower values around some of the hazard sites, sometimes increasing at greater distances, probably indicate a dissociation between homes and factories. Many large installations are situated in rural areas, or at coastal or estuarine sites, or grouped together in industrial estates so that there are few res-

idential PCs nearby. Oil refineries, nuclear power stations, and airfields, for example, virtually exclude human habitation within a km or more of their centres. Other hazards with evidence of local industrial-residential competition included chemical works, oil distribution terminals, crematoria, and harbours. However, not all hazard types exhibited these residential exclusions, notably the less segregated installations labelled as "safe sites".

Factories closely integrated with residential areas might show very local competition for space within individual PCs. This could result in artificially low densities of cancers per 1000 PCs. This was confirmed directly. Figure 1 shows SDRs surrounding the premises of 36 paper manufacturers, 21 soap manufacturers, 35 cotton spinners and weavers, 34 brewers, 31 mail order firms, and 38 furniture makers – the 195 locations aggregated as the "safe sites" of table 1. In these instances we know that the low short range case densities per 1000 PCs did not stem from low PC densities per km². Together with their later regression towards the mean regional values, these findings most probably indicate short range domestic industrial competition for space within PCs adjacent to or containing a factory.

The general regional relationship between low PC densities per km² and low case densities per 1000 PCs, the latter probably mediated through an excess of low DP PCs, also accords with these short range findings; but there was no way in which this could be specifically confirmed within the available data. Formal exclusion of a demographic artefact capable of mimicking source related case concentrations was not possible. However, this artefact remains entirely hypothetical, without the support of prior evidence; and despite these searches of our own data, no positive evidence in its favour was found. Indeed, the balance of evidence indicates an artefact capable only of reducing factory-adjacent cancer density estimates.

PETROLEUM BASED INDUSTRIES

Figure 2 shows the pattern of SDRs in separate bands around 15 major oil refineries, 17 major oil storage areas, and 168 lesser oil distribution terminals and storage sites – mostly rail to road or rail to factory transfer sidings or factory depots. Short range SDRs were irregular through small numbers; but for the refineries and major storage sites there was a clear SDR excess between 1 and 5 km (see table 2 for the aggregated results). Figure 2 also illustrates results for 45 manufacturers of industrial chemicals, excluding the 10 classified as petrochemical factories. Petrochemical factories showed a non-significant excess and the others showed none at all; nor did the six major UK benzene refineries, plants which at this time produced about a million tonnes per year. Critics of our previous paper thought that our results supported a link between benzene and leukaemia,⁶⁷ but we found no direct evidence on this, then or now.

Table 2 Proximities of death addresses to point hazards

Hazard type	No of sites	Distance (km)	No of cases	No expected	SDR*	χ^2	P
Oil refineries	15	1.0-5.0	329	259.9	1.27	18.34	<0.001
Oil farms	17	1.0-5.0	735	621.5	1.18	20.74	<0.001
Oil distribution terminals	168	1.0-5.0	3752	3411.0	1.10	34.08	<0.001
Bitumen products	22	0.0-5.0	1664	1512.8	1.10	15.11	<0.001
Detergent makers	33	0.0-3.0	1200	1039.3	1.23	24.85	<0.001
Plastics formers	28	0.0-2.0	283	249.2	1.14	4.59	<0.05
Car makers	31	0.0-5.0	2031	1789.2	1.14	32.70	<0.001
Coach builders	32	0.0-5.0	2025	1840.1	1.10	18.48	<0.001
Solvent makers	32	0.0-3.0	921	785.2	1.17	23.48	<0.001
Spray painting	33	0.0-5.0	2049	1846.7	1.11	22.16	<0.001
Galvanizers	69	0.0-5.0	6580	5922.1	1.11	73.09	<0.001
Paint makers	28	0.0-5.0	3679	3467.7	1.06	12.88	<0.01
Rubber manufacturers	97	0.0-5.0	5595	5273.2	1.06	19.65	<0.001
Varnish makers	13	0.0-2.0	337	286.3	1.18	8.97	<0.01
Fibreglass fabrication	37	0.0-3.0	1331	1201.2	1.11	14.03	<0.001
Adhesives makers	29	0.0-3.0	798	688.1	1.16	17.57	<0.001
Powder coating	26	0.0-5.0	784	682.4	1.15	15.11	<0.001
Steel works	44	0.0-5.0	1829	1739.5	1.05	4.61	<0.05
Aluminium casting	24	0.0-2.0	354	310.1	1.14	6.20	<0.01
Zinc casting	43	0.0-3.0	1523	1389.9	1.10	12.74	<0.001
Car batteries	14	0.0-3.0	459	363.0	1.26	25.37	<0.001
Power stations	69	0.0-3.0	1039	881.4	1.18	28.18	<0.001
Cement works	12	0.0-3.0	209	158.8	1.32	15.86	<0.001
Crematoria	64	0.0-1.0	302	214.6	1.41	35.60	<0.001
Crematoria	64	0.0-3.0	2167	1830.9	1.18	61.70	<0.001
Brickworks	27	0.0-2.0	142	117.6	1.21	5.08	<0.02
Airfields	192	0.0-3.0	876	743.6	1.18	23.57	<0.001
Rail yards	106	0.0-1.0	409	362.5	1.13	5.97	<0.02
Harbours	82	0.0-2.0	688	585.6	1.17	17.91	<0.001

* SDR (standardised density ratio) in this table accumulates individual bands shown in figures across the ranges given in column 3. Most sources gave significant results at several boundaries – exemplified here by the crematoria. For most hazards the table displays SDRs at 2.0, 3.0, or 5.0 km, usually selecting the maximum O/E (observed/expected) ratio. Significance levels were sometimes greater at other boundaries.

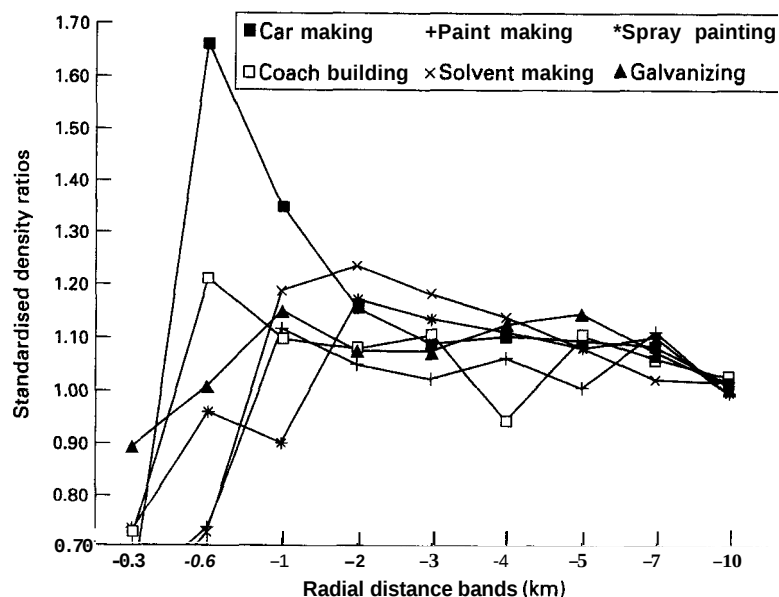


Figure 3 Automobile manufacture.

AUTOMOBILE INDUSTRY

It has been claimed¹⁵ that a greater quantity of volatile organic compounds is discharged when painting a car than is emitted from its exhaust during the whole of its road career. The 31 car factories examined (fig 3) showed a clear excess of leukaemias and cancers, which was greatest at 0.3–0.6 km, but extended up to 5 km. Coach building and body repair firms (32 sites) had a less powerful but significant effect over the same range.

Non-car factories pursuing related trades, and using similar organic materials, were also examined. Among them were 69 galvanizing plants – which use solvents for metal cleaning – and 33 spray painting contractors. They

showed pronounced excesses at ranges beyond the very shortest. Paint manufacture (28 sites) had a less powerful effect. These three, together with car manufacture and coach building, comprise the 193 sites whose adjacent PC densities are described in table 1. Brake manufacture (13 sites) and tyre making (23 sites) showed no effects. Other forms of rubber manufacture showed a small relative excess within 5 km (1.06) but none at shorter ranges (0–3 km). Among other solvent based processes examined, there were excesses in the vicinities of varnish makers, fibreglass fabricators, adhesives makers, and factories undertaking powder coating of metals. Despite the use of solvent based metal cleaning, 71 electroplating factories exhibited no apparent effect. Factories (22 sites) making halogenated hydrocarbons – chlorinated and fluorinated – likewise had no apparent effect, but 32 other solvent manufacturers showed highly significant cumulative excesses ($p < 0.001$) up to 5 km. After some short range irregularities, significant effects were also found for 22 makers of bitumen based products, 33 detergent factories, and 28 plastics factories.

METAL CASTING AND REFINING

The associations with car manufacture and with galvanizing suggest alternative non-solvent exposures: namely to metal casting, metal forming and welding. Examination revealed (fig 4 and table 2) short range risks associated with aluminium, zinc, and iron and steel casting. Steelworks showed a moderate effect. Firms offering contract welding construction showed nothing, but the addresses are probably remote from their on-site work locations. Casting and refining of lead (22 sites) showed no excesses, although car battery factories exhibited a strong effect. Lead casting is a relatively low temperature process and the battery making effect may be related to the manufacture of battery casings, plastics forming or solvent usage, rather than lead processing.

“Hot process” associations were not limited to metal forming. After some short range irregularities, powerful medium range excesses were evident for power stations, crematoria and cement works. Cement rail terminals had no effect. Brickworks showed a local effect but potteries (48 sites) and gasworks (98 sites) did not. The contrast between brick and pottery manufacture may be an effect of the relative scales of production, and the absence of an apparent hazard from gasworks may be explained in terms of local population exclusions or perhaps in terms of their history. In 1965/6 20 million tonnes of coal and oil were used in the production of gas, but by 1970/1 this was down to 8 million; and by 1976 to 0.2 million, as gas manufacture finally vanished.¹⁶ Even before 1970, and still in the era of gas manufacture, many local gasworks were closed or reduced to simple storage sites following the introduction of high pressure coal and oil gasification processes. These new processes, and the import of liquefied natural gas, led to the

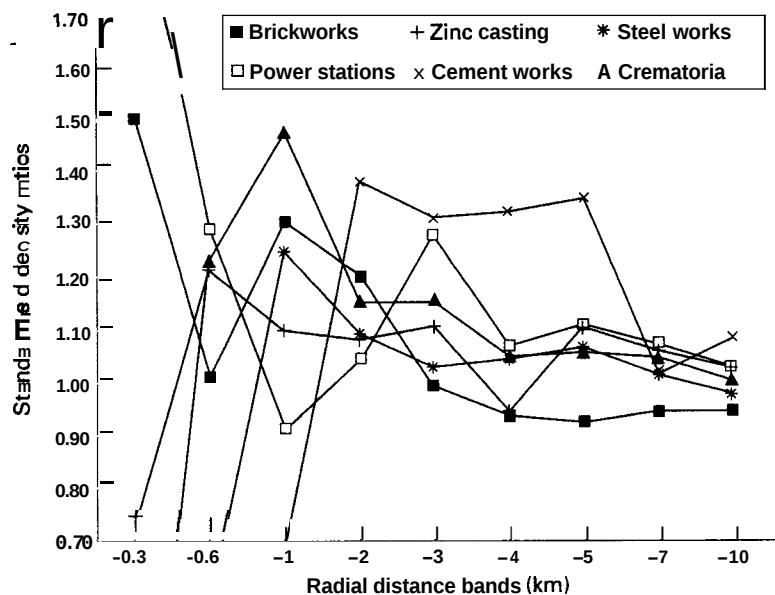


Figure 4 *Kills and furnaces.*

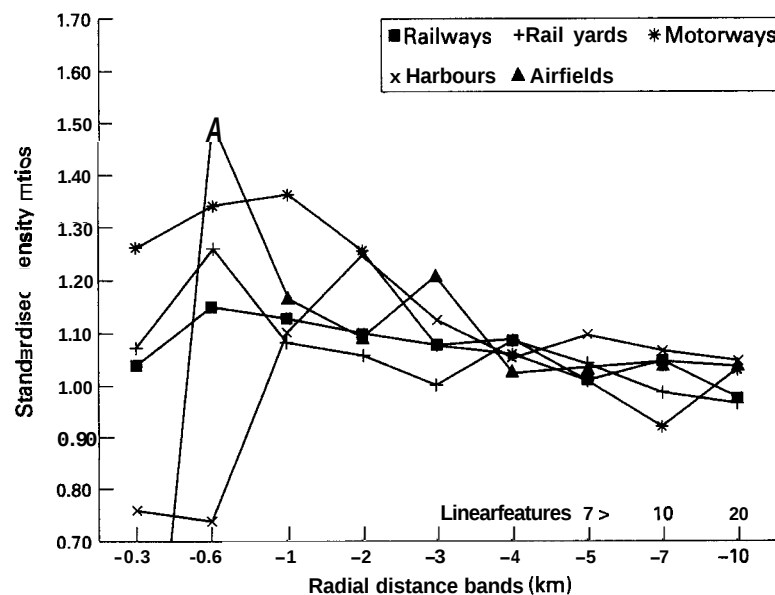


Figure 5 *Transport facilities.*

construction of high pressure long distance pipelines and to centralised production.

TRANSPORT ROUTES

Motorways

Distances were measured between all leukaemia/cancer bearing PCs and the nearest motorway (outside and excluding the M25) and a frequency distribution compiled. This was repeated for 10000 randomly selected PCs and the result used to calculate regionally standardised expected values. SDRs were calculated for an extended set of parallel distance bands up to and beyond 20 km. There was a significant excess of cases within 4.0 km of a motorway, and a significant deficit beyond 4.0 km. The calculation was repeated without the use of local standardisation, with the same result.

Railways

The railway network was examined in the same way. The pooled observed and expected numbers across all 11 rail regions showed a strong and significant excess within 2 (or 3 or 4) km of a rail line and a significant deficit beyond 4 km. The distance distributions differed in different regions but there were significant SDR excesses in at least one sub-4 km band in 9 of the 11 regions. The main exception was in industrial Lancashire and north Cheshire, which showed no difference. Here, a very high proportion of control PCs were within 2 km of a railway and thus afforded no effective basis for comparing near and far distances. Railway goods yards showed a significant short range SDR of 1.13. Harbours, most of them serviced by railway lines, and heavily laden with other industrial sources, showed an SDR of 1.17 within 2.0 km.

Distance related SDRs within 4 km of railways and motorways were readjusted using pooled data beyond 4.0 km as the reference level, and these adjusted relative risks are shown (on an extended scale) in figure 5. As with many industrial sites, as well as with harbours and airfields (also in fig 5), the railway risk was lower at 0–0.3 km than at rather greater ranges, probably from residential-industrial competition for space in railway adjacent PCs. For the motorways this was less pronounced. The rail and motorway analyses were repeated for the 9411 leukaemia registrations, with the same results.

OTHER INDUSTRIES

A wide range of industrial locations gave negative results. In addition to those already mentioned they included agricultural fertilizer rail terminals, MOD rail sidings, TV transmitters, cake and biscuit bakers, dry battery makers, magnetic tape makers, nuclear plants, PVC compound manufacturers, and makers of wood preservatives.

AIRBORNE DRIFT

Easting and northing biases of cases, relative to expected cases, were measured in all analyses. Nothing was found. For some, the calculations were repeated with specific drifts added to the coordinates of the hazards, to represent the effects of the prevailing wind. The results were compared with those of zero and counter prevailing wind drifts. No evidence of any drift was found and centrifugal risk gradients generally diminished with increasing imposed biases, in any direction. In the case of motorways, for example, a “dummy” linear hazard displaced 8 km N and 8 km E from the actual network annulled the cancer proximity pattern.

BIRTH ADDRESS OR DEATH ADDRESS?

This question was tested first against a composite list of those “hot process” sites whose individual effects had already been demonstrated. It included steelworks (44 sites), power stations (69 sites), cement works (12

sites), crematoria (64 sites), aluminium casting works (24 sites), and zinc casting works (43 sites). Proximities were measured against the following:

- The full list of death addresses,
- o The subset known not to have changed their PCs between birth and death,
- o The birth addresses of those who moved, and
- The death addresses of those who moved.

The full death address list showed an observed/expected (*OE*) ratio of **1.15** within 3 km. Non-movers showed a ratio of **1.19**. The birth addresses of movers also showed **1.19** and the death addresses of movers a ratio of **1.15**. All these excesses were highly significant ($p < 0.001$), suggesting that while early exposures were more important than later ones, they could also be effective at later ages. This corresponds with our earlier findings in relation to PC sharing pairs.² However, detailed examination revealed that many of those who had changed their PCs had moved only a short distance. The mean distance was **35.6 km**, but the distribution was highly skewed and **49.9 %** of this group had moved less than 5.0 km and **65.1 %** less than 10.0 km.

The examination was therefore repeated among **3213** children who had moved more than **10 km**, and where both addresses were accurately coded. Within 3.0 km of the hazard closest to either address, the birth places showed an *OE* a ratio of **1.09 (572/526.9)**. The ratio for the death addresses around the same sources was only **0.80 (420/526.9)**. Each was significant, as was the difference between the numerators (set on a common denominator). Their contrary directions reflect the imposed separation, through selection, of the two addresses; if one address is close to a particular hazard then the other is necessarily some distance away. Similar results were obtained for other hazard types and for different migration distances. Among children who had moved more than 5.0 km and where at least one address was within **10 km** of a galvanising plant, we found **186** births but only 83 deaths within 3 km of the nearest source.

We conclude that early exposure is the more important. Cancer initiations in migrating children may indeed be limited to exposures occurring within a short period before or after birth. The hazard proximities of the death addresses, as well as of birth addresses, in the total data set could reflect only their strong geographical associations.

Discussion

The logistical and methodological difficulties of conducting intimate proximity analyses on so large a scale are reflected in the paucity of comparable reported studies. Apart from several proximity studies of childhood cancers and leukaemias around nuclear installations,¹⁷⁻²⁰ we found little which could be related to the findings described here. Several large scale geographical-ecological investigations have related socioeconomic factors and population

densities to childhood leukaemia risks, but within much broader zones. Neither the relative distance methods developed for investigating particular nuclear sites¹⁹⁻²¹ nor the assembly of enumerated civil populations into, for example, 10 km squares or 25 km radius circles,^{17,22} met the technical needs of the present search.

The continuing demographic difficulties of a PC based method are clear, and a totally unambiguous demonstration of proximity clustering has still not been possible, but the present investigation stands on less error prone foundations than did the previous one.⁶ First, our recent analyses have effectively dispelled caveats about the reality of short range case clustering and the existence of geographically localised hazards is not now in doubt. Proximity studies are no longer concerned with this issue and can be directed solely at asking what those hazards might be. Second, some of the hazard associations match those previously postulated. The prior hypothesis of a petroleum associated leukaemia risk seems to be confirmed in a very large and independent set of disease data, and against a wider range of environmental sources. Third, the new analysis bases its estimates of geographical population distributions upon much larger numbers of PCs.

The main remaining source of ambiguity is a hypothetical systematic excess of children at risk in PCs close to suspect sources. A local excess of high DP PCs or of high child DPs could create a false apparent excess of cancers. Broad areal density standardisation (as used) provides no protection against hidden, small scale demographic structuring of this kind. Nor could the use of available population data, such as EDs or electoral wards, be expected to resolve this problem. However, our own data revealed low rather than high areal PC densities adjacent to many putative hazards. There were clear signs of competitive exclusion of residential PCs from the vicinities of very large plants: and of houses from those residential PCs which were very close to factories. Notably, low case densities were found within very short distances of non-volatile organic compound, non-combustion sites. Local demographic variations of these kinds could readily mask a genuine short range hazard related risk or create false appearances of a low risk; but they could not easily create a false excess. The major risk mimicking potential artefact therefore remains entirely hypothetical. There was never any prior evidence to show that it existed and our own indirect evidence now suggests that it does not.

Further reassurance against false positive demographic artefacts appeared in analyses of cases who had migrated between birth and death. The highly asymmetric ratios of **572/420** birth/death addresses within 3 km of a "hot process" source among children who had migrated 10 km or more and a **186/83** birth/death ratio within 3 km of galvanizing plants among children who had migrated 5.0 km or more are both independent of these demographic contingencies. These results also suggest a powerful general method for investigating age sensitive exposures where there are no demographic data available at all. We shall

describe its properties and develop its usage in a later study.

Effective apparent exposures were concentrated around two main kinds of installation namely: (a) producers, refiners, distributors, and industrial users of petroleum fuels and volatile petroleum products, and (b) manufacturing processes using high temperature furnaces, kilns, and combustion chambers. Some sources, notably internal combustion engines and oil fired furnaces, met both criteria. These specific associations and their absence around other sources were too coherent to suggest a demographic artefact. The few inconsistencies were probably due to different scales of production or restricted periods of operation. There was little to suggest systematic confounding from adjacent sitings of different processes. With few exceptions (eg pottery manufacture) the main factory based hazards were distributed independently across many regions. The rail and motorway associations were relatively ambiguous and could more readily be attributed to adjacent industrial hazards, but even here the most obvious direct common factor would be the use of petroleum fuel, especially diesel. Diesel is relatively non-volatile and it is difficult to envisage a spillage based or access based mechanism affecting both sources, so if fuel combustion is the common factor the most eligible materials must be blow-by fumes from crankcases and substances emitted from exhausts.

However, combustion fumes could not account for all the findings and there was no indication of any other class of effluent capable of explaining the full range of effects. None of the carcinogenic industrial species listed in the IARC monographs would easily serve as a common **determinant**.²³ The specific role of benzene, much discussed in the literature and in the media, was not supported by **an** examination of the major benzene manufacturing sites. Its reputed leukaemia toxicity may have led to special containment measures, but it would be surprising if none had escaped.

Rather than one specific effector, we must therefore consider the possibility that there might be several – from among the aromatics, other volatile organic substances, arsenic, metal fumes, soot and smoke compounds, radioactive elements, carbon monoxide, oxides of nitrogen and sulphur, and others. Alternatively, chemical species from different hazard types are known to combine in the atmosphere to form secondary pollutants.²⁴ The suspect primary substances are oxides of nitrogen and volatile organic substances, discharged singly or in combination. The secondary pollutants include ozone, aldehydes, ketones, nitrogenated organics, nitrates, hydrogen peroxide, ionised oxygen, and many others.

Alternative modes of access to such substances include the following:

- o Diffusing gases and volatiles reaching the children or their pregnant mothers directly,
- o Mediation through parental/occupational germ-cell injury,

- o Occupational/domestic contamination via clothing or through other means.

The last is feasible only for a limited range of sources (eg car factories or paint spraying contractors) and it is unlikely that sufficient numbers of parents were employed in oil storage installations and crematoria to explain the excess cancers in their vicinities. This is also against an exclusively parental occupational germ-cell or fetal exposure mechanism. For a comprehensive single explanation we must favour direct exposure of pregnant women or young children to airborne substances diffusing into the surrounding environment.

The apparent absence of tumour type specificities for different sources possibly reflects **an** intimate *mix* of several potent substances with different tissue specificities or a single common effector capable of reaching many different tissues and invoking a polyvalent response, as in the case of prenatal medical irradiation.⁹ Alternatively, it might reflect **only** the limited geographical resolution of our analyses. The O/E ratios are all modest and hazard adjacent type specific excesses could easily be buried among cases unrelated to the source in question.

Finally, we have to ask whether the moderately increased densities of cases around these hazards also explain the previously reported **high** local concentrations represented by same PC and same ED pairs.⁹ The first reached out as far as 3 km and sometimes 5 km, while the close pairs were detected within diameters of about 300 m. They do not necessarily stem from a common mechanism, although **two** considerations suggest that they might. First, identified sources must vary in the toxicities or containment of their emissions, and case densities surrounding the more dangerous could be nearer to those within the small clusters. Second, populations within **high** exposure zones are themselves so patchy that any resulting cancers are necessarily concentrated within small sub-areas.

We can not at present resolve these issues, and patterns of exposure could be far more intricate than our present results have indicated. Combustion based and volatile organic substances must be emitted by many less obvious sources at lower elevations, diffusing over shorter ranges, and leading to more concentrated effects than those surrounding major installations. Candidates include domestic and commercial heating systems, oil storage bunkers, oil delivery spillages, garage blocks, small workshops, bus stations, school or hospital chimneys, municipal incinerators, petrol stations, and many others. There are also many close clustered exposures to non-effluent hazards, each capable of reaching cellular targets and damaging the genetic material, or immunological processes, or the mechanism of cell division. They include domestic radon and gamma radiation, medical radiation, ambient electromagnetic fields, infections of mother or child and prenatal drug exposures. All of them have been implicated although none offers a comprehensive and exclusive explanation.^{9 22 25-30} To solve these questions we need

an extended and better differentiated record of exposures, and more discriminating modes of examination.

Appendix

CALCULATION OF STANDARDISED DENSITY

RATIOS (SDR)

Consider a set of h hazards of similar type, H .

The area surrounding each hazard consists of concentric annuli, A , containing O_{hr} observed cases and P_{hr} PCs.

Each hazard is located in one of 56 rectangular zones Z_z , in which k cases and n PCs give a zonal case density $D_z = k_z/n_z$ and the expected number of cases in each of the A_m annuli is

$$E_{hr} = P_{hr} * D_z.$$

Summing observed and expected values for all hazards of a given type gives the SDR for each successive annulus around a hazard of type H as

$$SDR = \frac{\sum_h O_{hr}}{\sum_h E_{hr}}$$

The functions of regional standardisation are to calculate local expected numbers, weighted according to corresponding regional case densities; and to calculate ratios (SDR) between the aggregated observed and expected numbers. For restricted subsets of cases, such as specific tumours, or children who moved house, the regional case densities are first reduced in proportion with the size of the subset.

We thank Dr A M Stewart for access to the data of the Oxford Survey of Childhood Cancers. The postcoding of the OSCC records was part of a research programme supported by the Medical Research Council and by the Three Mile Island Public Health Fund (USA). EGK's expenses during the current work were defrayed through a Leverhulme Emeritus Research Fellowship.

- 1 Knox EG, Gilman EA. Leukaemia clusters in Great Britain. 2. Geographical concentrations. *J Epidemiol Community Health*. 1992;46:573-76.
- 2 Knox EG, Gilman EA. Spatial clustering of childhood cancers in Great Britain. *J Epidemiol Community Health*. 1996;50:313-19.
- 3 Gilman EA, Knox EG. Temporal-spatial distribution of childhood leukaemias and non-Hodgkin lymphomas in Great Britain. In: Draper GJ ed. *The geographical epidemiology of childhood leukaemia and non-Hodgkin lymphomas in Great Britain 1966-83*. Studies on medical and population subjects No53. London: HMSO, 1991.

- 4 Knox EG, Gilman EA. Leukaemia clusters in Great Britain. 1. Space-time interactions. *J Epidemiol Community Health*. 1992;46:566-72.
- 5 Gilman EA, Knox EG. Childhood cancers: space-time distribution in Britain. *J Epidemiol Community Health*. 1995;49:158-63.
- 6 Knox EG. Leukaemia clusters in childhood: geographical analysis in Britain. *J Epidemiol Community Health*. 1994;48:369-76.
- 7 Bithell JF, Draper GJ. Apparent association between benzene and childhood leukaemia: doubts concerning a report by Knox. *J Epidemiol Community Health*. 1995;49:437-39.
- 8 Stewart AM, Webb J, Hewitt D. A survey of childhood malignancies. *BMJ* 1958;1:1495-508.
- 9 Knox EG, Stewart AM, Kneale GW, Gilman EA. Prenatal irradiation and childhood cancer. *Br Radiol Prot* 1987;7:177-89.
- 10 Stiller CA. Population based survival rates for childhood cancer in Britain. *BMJ* 1995;309:1612-26.
- 11 *Kompass: CBI-UK*. London: Reed Information Services, 1993.
- 12 *Kelly's*. London: Reed Information Services, 1994.
- 13 Alexander FE. Investigations of localised spatial clustering, and extra-Poisson variation. In: Draper GJ ed. *The geographical epidemiology of childhood leukaemia and non-Hodgkin lymphomas in Great Britain 1966-83*. Studies on medical and population subjects No53. London: HMSO, 1991.
- 14 Muirhead CR. Childhood leukaemia in metropolitan regions in the United States: a possible relation to population density? *Cancer Causes and Control*. 1995;6:383-8.
- 15 Gibbs WW. An ounce of prevention. *Scientific American*. November 1994;271:84.
- 16 Williams TI. A history of the British gas industry. Oxford: Oxford University Press, 1981.
- 17 Cook-Mozzaffari PJ, Ashwood FL, Vicent T, Forman D, Alderson M. *Cancer incidence in the vicinity of nuclear installations, England and Wales 1959-80*. Studies on medical and population subjects No51. HMSO. London, 1987.
- 18 Cook-Mozzaffari PJ, Darby SC, Doll R, et al. Geographical variation in mortality from leukaemia and other cancers in England and Wales in relation to proximity to nuclear installations, 1967-78. *Br J Cancer* 1989;59:476-85.
- 19 Bithell JF, Stone RA. On statistical methods for analysing the geographical distribution of cancer cases near nuclear installations. *J Epidemiol Community Health*. 1989;43:79-85.
- 20 Bithell JF, Dutton SJ, Draper GJ, Neary NM. Distribution of childhood leukaemias and non-Hodgkin Lymphomas near nuclear installations in England and Wales. *BMJ* 1994;309:501-5.
- 21 Bithell JF. Statistical methods for analysing point-source exposures. In: Elliott P, Cuzick J, English D, Stern R eds. *Geographical and environmental epidemiology: methods for small-area studies*. Oxford: Oxford University Press, 1992: 221-30.
- 22 Knox EG, Stewart AM, Gilman EA, Kneale GW. Background radiation and childhood cancers. *J Radiol Prot* 1988;8:9-18.
- 23 International Agency for Research on Cancer. *WHO monographs on the evaluation of carcinogenic risks to humans*. In particular, Vol 60. *Some industrial chemicals*. 1994. Lyon: IARC.
- 24 Kelly TJ, Mukund R, Spicer CW, Pollack AJ. Concentrations and transformations of hazardous air pollutants. *Environ Sci Technol* 1994;28:378-87.
- 25 Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukaemia. *Leukaemia*. 1988;2:120-5.
- 26 Kinlen LJ. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain. *Lancet* 1988;2:1323-27.
- 27 Kinlen LJ, Clarke K, Hudson C. Evidence from population mixing in British new towns 1946-85 of an infective basis for childhood leukaemia. *Lancet* 1990;336:577-82.
- 28 Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. *Br J Cancer* 1995;71:1-5.
- 29 Kinlen LJ, John SM. Wartime evacuation and mortality from childhood leukaemia in England and Wales in 1945-9. *BMJ* 1994;309:1197-202.
- 30 Coleman M, Beral V. A review of epidemiological studies of the health effects of living near or working with electricity generation equipment. *Int J Epidemiol* 1988;17:1-13.