Migration patterns of children with cancer in Britain

E G Knox, E A Gilman

Abstract

Objectives—To investigate the early migration patterns of children who later developed cancer. To test a prior hypothesis that some cancers are initiated by early exposures to toxic atmospheric pollutants from point sources.

Design—Address changes in children dying from cancer are examined in relation to potentially hazardous sites of several different types. The relative proximities of birth addresses and death addresses to these sites, are compared. The approach is based upon the premise that a local exposure, effective only at an early age, must be preferentially linked with an early address.

Setting and Subjects—Records of 22 458 children dying from leukaemia or other cancer under the age of 16 years in Great Britain between 1953 and 1980: including 9224 who moved house between birth and death. The migration analysis was based upon birth and death addresses, converted first to postcodes and thence to map coordinates. The geographical locations of potentially toxic industrial sites were obtained through direct map searches and from commercial directories.

Results—Systematic asymmetries were found between measured distances from birth and death addresses to sources emitting volatile organic compounds, or using large scale combustion processes. The children had more often moved away from these hazards than towards them. Many of the sources had already been identified as hazardous using other methods. There was also a birth association with areas of dense habitation; possibly because of unidentified toxic sources contained within them. All forms of cancer were involved although some effluents were associated preferentially with specific types.

Conclusions—The main findings of an earlier study, based upon a different and independent method, were confirmed. Proximities to several types of industrial source, around the time of birth, were followed by a raised risk of childhood cancer. Combustion products and volatile organic compounds were especially implicated. Within the 16 year limit of the study, the increased risk did not decay with advancing age. Low atmospheric concentrations of many carcinogenic substances suggest that the mother acts as a cumulative filter and passes them to the fetus across the placenta or in breast milk.

(*J Epidemiol Community Health* 1998;**52**:716–726)

Mill Cottage, Great Comberton, Pershore, Worcs WR10 3DU
E G Knox

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

Correspondence to:
Professor Knox.

Accepted for publication
1 June 1998

We previously showed (a) that childhood cancers and leukaemias in Great Britain tend to occur in small geographical clusters: (b) that they occur at increased densities around industrial sites using high volume, high temperature furnaces, or discharging volatile organic compounds (VOCs).[1-4] Furthermore, among children who moved house between birth and death, the birth addresses were closer to these hazards than were the death addresses, suggesting that the exposures were more effective shortly before or after birth than at later ages. This matches much other evidence that these diseases are initiated early, and probably prenatally.

The hazard proximity measurements were based upon uncertain population estimates, and the only available group of control children had been matched for residence against the death addresses of the cases themselves. They had been selected as non-migrants to the extent that they had been born in the same residential area; so they were not suitable for case-control comparisons of geographical locations. However, the birth-death contrasts suggested an alternative form of group comparison. This was based upon the simple premise that where a child had moved more than some minimum distance, it was not possible for both addresses to be close to a specific and age limited effective source. Suitable tests of site toxicity could then be based upon measuring the interaction between the hazard and the two addresses of the same child, rather than between a hazard and the addresses of a sick and a "control" child. Suitable controls could still have been desired but might prove less than essential in a situation where long term geographical clustering was already well established, where long term focal hazards could confidently be inferred, and where only their nature remained to be tested and confirmed.

This approach could offer advantages even in situations where local populations have been enumerated. Age specific population numbers within short radial distances of point hazards are seldom directly available and they usually demand error prone estimates. Hidden fine scale variations and inapparent child concentrations or depletions, relative to geometrically defined boundaries, can result in false attributions of excess risks or the masking of genuine risks. The demonstration of birth-death contrasts based upon exact locations, circumvents these particular artefacts.

Our objectives are (1) to examine the methodological aspects of "birth-death hazard interactions" in more detail than previously, and (2) to expand earlier searches for cancer initiating environmental exposures.

Methods

The case material was extracted from a file of all 22 458 deaths from cancer or leukaemia occurring before the 16th birthday in Great Britain between 1953 and 1980. All fatal tumours were included and classified into 11 main groups (lymphatic, myeloid, monocytic and unclassified/other leukaemias: lymphomas, nephroblastoma, CNS tumours, neuroblastoma, bone cancers, other solid cancers, and fatal benign tumours). Home addresses at death were recorded, and where parents were subsequently interviewed, the birth address was also obtained. Postcodes (PCs) were identified, and the map reference of the "first" address within the PC was extracted from the Central Postcode Directory. The data set has been described in detail elsewhere.[3][4] In 9224 cases (41.1%) the child was known to have moved at least 0.1 km between birth and death, this group providing the main basis for this study. Another 9328 children (41.5%) had not moved. We had no information on the previous movements of the remaining 3906 (17.4%) fatal cases. The dates of the house removals were not recorded.

Where addresses or PCs were considered inexact or unreliable—usually for birth addresses—the grade of uncertainty was recorded. The best recorded were located within the available precision of the PC Directory: in urban areas within about 0.1 km of the actual address. Almost all of the 9328 non-migrating cancers, and 4385 of the 9224 migrating cancers (both addresses), were located to this level of accuracy. Using less stringent criteria, another 4818 migratory birth addresses were probably accurate within about 1.0 km. The remaining 21 were too uncertain for use in any critical context. Only the most accurate of the map reference pairs were used for close range migration analyses and the less accurate were used mainly for those wider scale examinations where precision was less critical. They were also used to double check shorter range studies to test the possibility that artefacts might have arisen from geographically biased "selection by certainty".

The sites of putative hazards were identified through direct map searches and through reference to general business directories and specialised lists. Most of them have been listed elsewhere,[4] but we carried out a number of additional searches (see tables and appendix 1). As before, the administrative headquarters of large firms were avoided in favour of active manufacturing sites; while sites and processes established or relocated only in recent years were not used. We included many previously active sites that are now closed, such as older power stations and local gasworks.

Two modes of analysis were examined and tested. The first sought out the birth and death locations of children around the sites of the hazards. The second looked for hazard sites around each of the two addresses. The difference is more fundamental than it sounds (see appendix 2 for details). Both methods compared the distributions of birth to hazard and death to hazard distances, and resolved them into northerly and easterly components, but the second ("case centred") method used only the nearest hazard. It measured and displayed individual (and mean) migration distances and directions relative to the nearest source; and it expressed centrifugal migration as a simple ratio between numbers of children migrating away from or towards the hazard.

Mean distances and differences were tested against calculated standard errors, and the outward/inward ratios through χ^2 tests. p Values calculated for many separate examinations within a systematic search for asymmetries, are not to be interpreted as tests of precise hypotheses; nor indeed is there any valid formal approach to the "significance" of such findings. The results should be treated only as informal guidance on the extent of sample variability.

The case centred method also generated simple and effective visual displays. Where a single source was within a limited distance of both addresses, the line joining the two points could be drawn on a circular map of the surrounding area. This provided a visual indicator of the magnitude and direction of the movements relative to the hazard. With several sources registered upon a common centre (see figures), and with birth addresses distinguished from death addresses, these lines served as "compass needles" attracted to the hazard: also capable of demonstrating its highly localised (or more diffuse) nature. Diffuse centres of attraction indicate area sources rather than point sources, or geographical confounding between adjacent point sources.

The case centred method permitted the prior selection of the most informative address pairs; for example excluding those who had moved only short absolute or short radial distances. It also permitted specific adaptation to discharge points whose elevations or relative volumes implied a wider or narrower dispersion. Broad field examinations first compared total numbers of address couplets pointing towards or away from a hazard. However, more critical examinations then used only those address pairs lying on opposite sides of an intermediate boundary: for example, inside and outside a 3 km or a 1 km limit. Two intermediate limits could be set to exclude an intervening neutral zone; for example, requiring two addresses on opposite sides of a 2 km to 5 km neutral zone. Processes discharging through roof top ventilators or at ground level were tested across narrow intermediate boundaries, against expectations that even short distance outward migrations might result in reduced exposures. On the other hand, waste products discharged through tall chimneys create shallow and sometimes complex radial exposure gradients so that only longer migrations would be able detect them. We refer to address pair selections using one or two intermediate boundaries as the zonal contrast approach.

The operational flexibility of the case centred approach, its statistically efficient exploitation of paired addresses, and its exclusion of all but the most intimately related sources, readily accounted for the superior sensitivity and

Table 1 Distributions of minimum distances between object locations and the birth and death addresses of migrant cancer cases

(A) Minimum distances between object set of 10 000 random PCs and 4173 migrant cancer cases*

Minimum distance (km)	Number of addresses		Ratio	p Value
	Object to birth	Object to death		
≤0.6	1609	1457	1.10	<0.01
−1.0	1129	1100	1.03	
−2.0	855	961	0.89	
−3.0	250	265	0.94	
−5.0	188	255	0.74	
−10.0	110	112	0.98	
>10.0	32	23	1.39	
Total	4173	4173		
Mean	1.261	1.328	dif/SEM=2.51	<0.02

Relative distances from nearest object site to:

	Birth address nearest	Death address nearest	χ^2	p Value
All cases*	2225	1948	18.25	<0.001
Short migrations†	1211	1138	2.21	NS
Long migrations‡	1013	810	22.38	<0.001

(B) Minimum distances between object set of 9297 non-migrant cancers and 4057 migrant cancer cases*

Minimum distance (km)	Number of addresses		Ratio	p Value
	Object to birth	Object to death		
≤0.6	1718	1398	1.23	<0.001
−1.0	987	1030	0.96	
−2.0	668	877	0.76	
−3.0	272	258	1.05	
−5.0	217	257	0.84	
−10.0	152	191	0.80	
>10.0	43	46	0.93	
Total	4057	4057		
Mean	1.346	1.502	dif/SEM=3.87	<0.001

Relative distances from nearest object site to:

	Birth address nearest	Death address nearest	χ^2	p Value
All cases*	2269	1788	56.79	<0.001
Short migrations†	1211	1014	17.27	<0.001
Long migrations‡	1058	774	43.72	<0.001

*All cases where both addresses are within 20 km of nearest object. †Subset migrating ≤5.0 km. ‡Subset migrating >5.0 km.

specificity that it displayed in use. The "nearest hazard" constraint and the use of tailored preselection processes presumably deleted many random relations and much geographical confounding between hazards of different kinds. For these reasons it was adopted as the preferred method and it is used exclusively in the account that follows.

Results

DEMOGRAPHIC RELATIONS

Both methods were first tested against a dummy object set of grid intersections, and this confirmed their technical neutrality. Case centred analyses were then performed against object sets of PC coordinates extracted randomly from the Central Directory. This was to provide a general background against which to examine specific nominated sources. For both smaller and larger samples of PCs (500, 1000, 5000, 10000, etc) we found that the birth addresses were closer to the random PCs than were the death addresses. This was at first surprising but the results were checked with great care; and repeat runs that inverted birth and death addresses in alternate cases abolished the birth-death differences. The finding was real.

Panel 1 of table 1 shows this asymmetry among those 4173 migrant cancer children whose PCs were both regarded as certain (4385) and where both addresses were also within 20 km of a "source". The minimum distance distributions differed, with a significantly greater mean value for the deaths than for the births: and a significant excess of very short (≤0.6 km) minimum distances for the births. The mean intra-pair death minus birth difference was significantly greater than zero: and the birth to object minimum was more often shorter than the pair matched death to object minimum. The examination was repeated with the less certain migrant addresses included, and with random object sets of different sizes. Mean distances varied with the size of the object set but the asymmetries were consistent. The tests were repeated for migrants dying in the first half of the survey period, and in the second half: and the results were again similar.

Random PCs are geographically concentrated in the same way as the population and in no sense do they represent random locations. The findings indicate only that the birth addresses of cancer migrants were located in areas with more PCs per sq km, and therefore with shorter mean inter-PC distances, than the areas to which they migrated. When sets of random PCs were subdivided into those whose adjacent PCs (in the full Directory) were ≤0.2 km apart, and those where the distance was greater than this, the birth-death asymmetry among cancer migrants was elicited only by the more crowded object set.

Several additional analyses confirmed that outward migrations were associated with high population densities, with a reverse pattern for low densities. Thus, the centroids of 105 country towns without rail connections and with populations between 2000 and 20 000, were assembled as an object set. The absent rail criterion was adopted because of earlier evidence of high case densities alongside railways, and because it excludes large concentrations of heavy industry. The intra-pair outward-inward migration ratio across an intermediate boundary set at 10.0 km showed a reversed movement (91:108), compared with the general pattern for random PC object sets (for example, 2225:1948 in table 1). That is, more children had moved towards these small town locations than away from them. The same was found for 44 major TV transmitters (110:142), again mainly in PC sparse rural areas. By contrast, when the locations of 204 non-league football grounds were used as a non-hazardous index of suburban and semi-urban environments, the identically measured outward:inward migration ratio of 636:552 resembled that for the random PCs. A set of 48 cathedrals, mostly in large cities, served as a non-hazardous indicator of the central urban environment. The migration ratio across a similar intermediate boundary (10 km), was 531:305.

The only factor that fundamentally changed the patterns of density associated outward migration near random PCs was the use of a narrower intermediate radius. Zonal contrasts of cancer migrations across boundaries within 2 km of PCs in different random object sets

Table 2   Short range asymmetries among birth-death address pairs. Crossings of intermediate boundaries around VOC sources*

(A) Separate sources

Hazard type	Sites	Births:x:deaths†	Birth-nearest: death-nearest	p Value (within-pair)
Car making/coachbuilding	63	160:3:128	172:120 (1.43)	<0.025
Spraypainting	33	56:2:31	128:81 (1.58)	<0.005
Galvanising	68	40:1:21	202:144 (1.40)	<0.005
Electroplating	74	202:3:165	234:162 (1.44)	<0.001
Powdercoating	26	34:2:21	62:51 (1.22)	NS
Fibreglass-form	37	75:3:41	83:50 (1.66)	<0.01
Rubberproducts	97	117:2:87	273:195 (1.40)	<0.001
Solvent/glues	65	165:3:136	388:302 (1.28)	<0.005
Paint manuf	28	118:3:91	121:92 (1.31)	<0.1
Varnish manuf	13	32:2:20	52:45 (1.16)	NS
Furnit manuf	38	76:2:46	150:96 (1.56)	<0.001

(B) Aggregate of above VOC sources and random PC controls. Use of different intermediate boundaries

	Sites	Intermediate boundary (x)	Birth-nearest: death-nearest	
Aggregate	542	nil‡	832:589 (1.41)	<0.001
Control	542	nil	840:815 (1.03)	NS
Aggregate	542	1	111:61 (1.82)	<0.001
Control	542	1	112:112 (1.00)	NS
Aggregate	542	2	171:113 (1.51)	<0.001
Control	542	2	192:167 (1.15)	NS

*Birth and death 1.0–6.0 km apart : both within 10.0 km of a hazard. †"x" is the intermediate boundary in km. 160:3:128 means that 160 births and 128 deaths were within 3 km of a hazard site. ‡"Nil" means that all pairs were counted without reference to any intermediate boundary. In paired couplet proximity comparisons any equal distances are counted in both classes.

generally failed to elicit these asymmetries. We return to this finding later.

DIRECT RISK INDICATORS
Examinations similar to those around the random PCs, were then performed around the addresses of 9297 non-migrant cancers, using identical acceptance criteria for the migrants: 4057 migrants from a total of 4385. The results are shown in panel 2 of table 1. Both panels show shorter and longer migrations (≤5 km, >5 km) separately. Object sites marked by the presence of a non-migrant cancer were more strongly associated with outward migration than were the random PCs; and especially for the shorter migrations.

The overall outward-inward migration asymmetry is represented again in a zonal con-
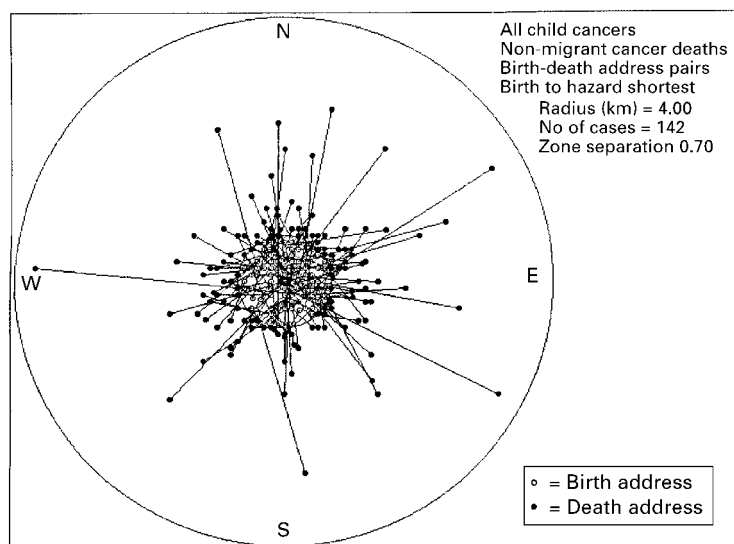

Figure 1   Movements of migrant cancers away from non-migrant cancer addresses (142 migrations from within 0.7 km of object site to the 0.7–4.0 km annulus).

trast format, but using different admission criteria, in figures 1 and 2. The centres of these circular maps represent the overlaid object sites, the outer circle represents the limits of admission to the examination, and the intermediate boundary is crossed by each admitted child. The couplets are shown between the birthplace, a hollow dot, and the deathplace, a solid one. Outward and inward migrations are shown in separate figures. Here, an intermediate boundary of only 0.7 km, within an outer limit of 4.0 km, showed a significant outward migration ratio of 142:108. This contrasted with a neutral value of 121:129 when tested at the same short range against the same number of random PCs.

The very local nature of these effects was confirmed by imposing randomly orientated fixed distance perturbations upon the non-migrant object-sites. Randomly orientated shifts of 0.7 km abolished the asymmetry, changing the ratio from 142:108 to 71:62. An alternative perturbation is shown in figures 3 and 4 where the same object sites were displaced by fixed amounts in a fixed direction: 0.7 km east and 0.7 km north. Here, and for other bearings as well, the former abundance of pointers was again reduced and the outward/inward asymmetry again abolished. The displaced pattern of points in these figures emphasises the very local—and now off-set—nature of the zones which generated these differential movements.

The same forms of examination were repeated using two other cancer site object sets, namely the birth addresses and the death addresses of the 4835 "address certain" migrant cases. These two object sets were tagged with their index numbers to avoid subsequent matches between migrant address pairs and their self derived object locations. Each migrating case was thus tested against the birth and death locations of every migrant except itself. The results for the birth locations of these migrants closely resembled those for the non-migrant sites with an outward-inward ratio of 2503:1907 (1.313). The respective numbers of birth and death addresses within 0.6 km of an object site were 1418:1003. The migration ratio in a zonal contrast examination similar to that of figures 1 and 2 was 160:116. All these short distance ratios were highly significant and far exceeded those for random PCs.

These studies were repeated for migrants dying in earlier and later parts of the survey, and for those dying at different ages; for example those dying before their third birthdays, those dying after their sixth birthdays, and those dying between 3 and 6 years. All gave similar results.

However, the use of the migrants' death addresses as the object set gave quite different results; none of the close range examinations showed a significant excess of outward migrations and some were orientated in the opposite direction. The overall outward-inward ratio was 2269:2195; the numbers of births and deaths within 0.6 km of an object site were 1177:1225; the comparable 0.7 km zonal
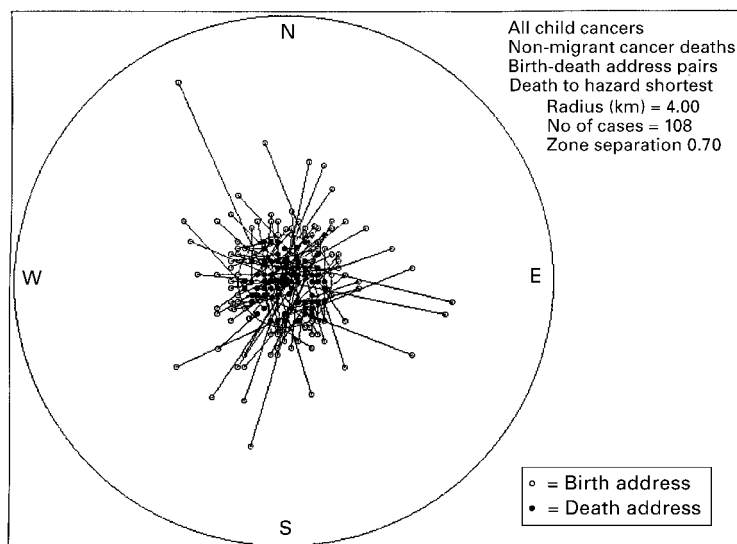
Figure 2   Movements of migrant cancers towards non-migrant cancer addresses (108 migrations to within 0.7 km of object site from the 0.7–4.0 km annulus).

contrast intra-pair ratio of figures 1 and 2 was 123:150.

Spatially confined migration asymmetries around sites marked by the birth of another affected child, differ so sharply from equivalent tests around random PCs and migrant death places, that local hazards must have been responsible for these early cancer initiations. The results confirm earlier evidence of spatial clustering, and the neutrality of the death locations confirms a movement from hazardous to less hazardous places. But these tests say nothing yet about the types of hazard responsible.

Interpretation of the less striking and less localised migration asymmetries surrounding the random PCs, is difficult. PC dense areas could be hazardous to the fetus or infant through associated industrial effluents, rail and road traffic, domestic and other chimneys, or exposure to infections, or the local idiosyncrasies of medical care; or the effect might be essentially demographic. Secular migrations
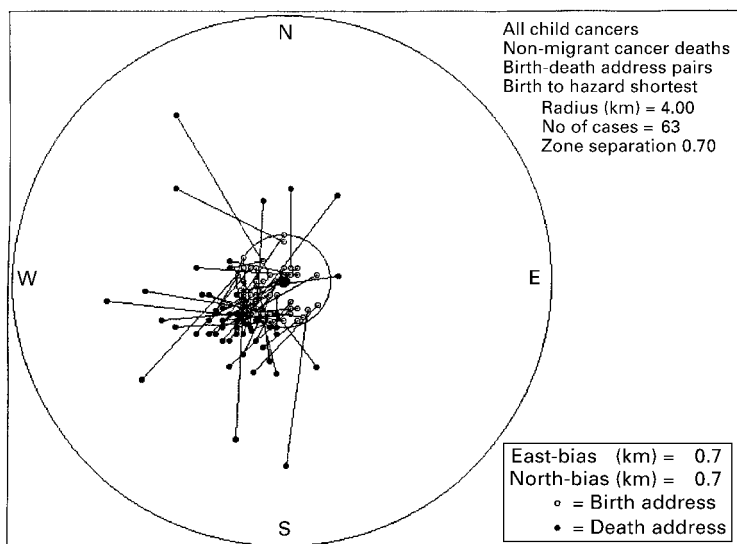

Figure 3   Movements of migrant cancers away from non-migrant cancer addresses, offset to north and east by 0.7 km (63 migrations compared with 142 in fig 1).

affecting the whole of the child population could have resulted from the clearance and replacement of dense sub-standard housing. Even within a constant housing stock, young families might preferentially have moved to less PC crowded neighbourhoods. We defer further interpretation until more evidence has been displayed, noting first that our studies of nominated industrial sources must each be compared with a random PC object set of similar size, tested in the same way.

SPECIFIC SHORT RANGE EXPOSURES

First examinations of nominated industrial sites were directed towards localised sources, with low elevation emissions, whose high adjacent case densities had already been demonstrated.[4] They included many factories emitting VOCs. Most were small to medium installations although some, such as car factories, were much larger. Effluents emitted near ground level create high local concentrations that taper rapidly with increasing distance, and they justify a search for shorter range, and therefore geographically more specific, migration effects. They also demand the use of the most accurate of the address specificatons. The first group of sites examined in this way, and the main results, are shown in table 2.

Panel 1 shows groups of object sites manufacturing or using VOCs. Many showed marked asymmetries in the numbers of births and deaths within a short critical distance (for example, 1 to 3 km) and/or high ratios of individual outward:inward migrations. These tabulations were limited to children migrating short distances (1.0–6.0 km) where both addresses were within 10.0 km of a hazard. The findings confirm our previous results based on geographical case densities[4] and add three other hazards that had not previously shown a risk: namely furniture and rubber manufacture, and electroplating. This probably reflects the improved sensitivity of the new method.

An aggregated VOC related object set was subsequently constructed from these sites. Results are given at the foot of table 2 and compared with a random PC object set of equivalent size, using several different criteria. The first comparison uses the paired addresses of all migrations confined within 10 km of a source. The others compare contrast zones separated by a 1.0 km and by a 2.0 km intermediate boundary. The migration asymmetries surrounding these sources were much greater than those around random PCs analysed in the same way. A graphical illustration of one zonal contrast result, for the reticuloendothelial (RE) malignancies, is given in figures 5 and 6.

Table 3 demonstrates similar short range effects around hazard types not prominent as VOC sources, but with previous evidence of adjacent cancer concentrations. Many emit effluents near ground level and they can be tested locally in the same way as the VOC group. They include detergent makers, harbours, railyards, zinc foundries and iron/steel foundries. These findings also implicate several
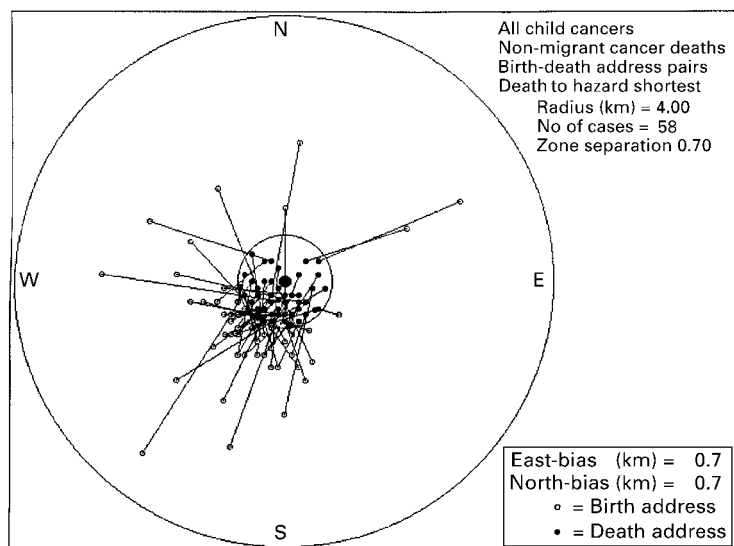
Figure 4 Movements of migrant cancers towards non-migrant cancer addresses, offset to north and east by 0.7 km (58 migrations compared with 108 in fig 1).

sources not previously showing high cancer concentrations; namely gasworks, soapworks, potteries and breweries: also hospitals, not previously tested. The last group consists of principal hospitals administered by the new (English) NHS Trusts. Large hospitals usually have incinerators, and their notorious and often low level emissions of black smoke long escaped prosecution through a legal immunity enjoyed by "crown property".

Similar tests confirmed the relative innocence of several processes that had not previously shown a risk. They included nuclear installations, mail order firms, brake manufacturers, lead casting plants, paper cotton and pesticide manufacturers, and bread and biscuit makers.

MEDIUM RANGE EFFECTS
The varied nature of the sites shown in table 3 did not justify their overall aggregation, but the combustion-based processes among them were combined for joint examination with several
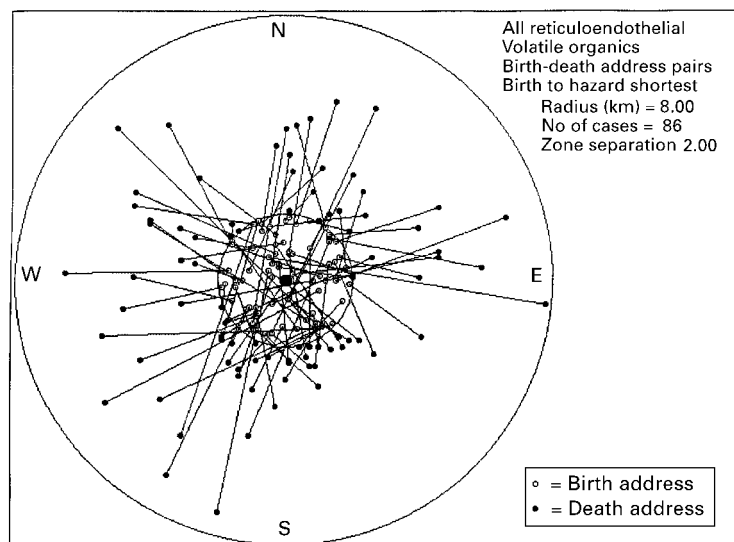

Figure 5 Migrations of cancers away from sites emitting VOCs (86 from within 2.0 km to the 2.0–8.0 km annulus).

KEY POINTS
- Childhood cancers occur in small geographical clusters spanning one to three postcodes.
- The ratio of cancer deaths to numbers of postcodes is raised near industrial sources that emit volatile organic compounds (VOCs) or that use large scale high temperature combustion processes.
- Among cancer children who move house between birth and death, their birth addresses are nearer to these hazards than are their death addresses. Leukaemias and other cancer types behave similarly.
- Many childhood cancers are probably initiated by fetal or early postnatal exposure to VOCs and to combustion products in the atmosphere, first absorbed and accumulated by the mother and transmitted through the placenta or in breast milk.
- In diseases with early origins and long latent intervals, interactions between radiating migration pathways and putative focal hazard sites, supply a powerful tool for the identification of toxic sources.

"tall chimney" combustion sources. Hospitals were studied separately. Emissions of such varied ranges and volumes, blanketing broad environments holding other hazards, demanded a "neutral zone" zonal contrast approach using two intermediate boundaries. Table 4 gives the main results. There were outward-inward migration excesses both for the hospitals and for the other combustion sources. Each was significant in its own right and differed significantly from comparable results obtained from a similar sized set of randomly selected PCs.

LONGER RANGE EFFECTS
Sources configured for very long range diffusion of their effluents raise formidable detection difficulties. The problems arise from the use of very tall chimneys and the use of enormous tonnages of fuels and volatiles; also from surrounding concentrations of other industries as well as roads and railways. There may also be few children within a short distance of very large remotely sited plants.

Oil refineries fall into this class. They are examined, together with major oil storage sites and lesser oil terminals, in table 5.

The widely diffused exposures from the larger sites justified the additional use of the less exact migrant address pairs (making 9203 altogether), omitting only the least satisfactory codings. Each of the oil installation classes showed statistically significant migration asymmetries across several alternative intermediate boundaries between 6.0 and 25.0 km. At these ranges, however, so did equivalent sized samples of random PCs.

As the examined ranges were narrowed, the asymmetries around the random PCs diminished, while those around the oil installations increased. This allowed more critical discrimination, except for the refinery examinations
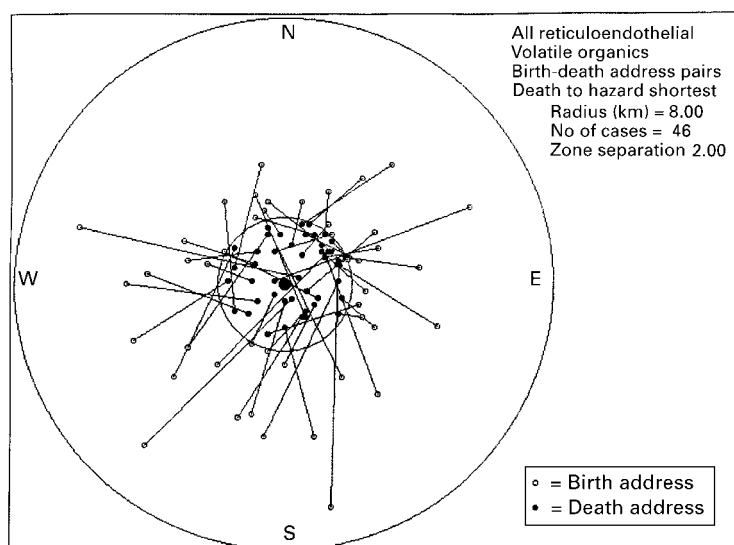
Figure 6   Movements of cancers towards sites emitting VOCs (46 to within 2.0 km from the 2.0–8.0 km annulus).

that now had too few nearby migrants. Short range circles around "any oil installation" showed a quite extreme outward:inward asymmetry; 242:82 for migrations across a neutral zone bounded by an inner radius of 1.0 km and an outer one of 2.0 km, compared with 210:209 for an equivalent random PC object set. The "certain" address pairs gave similar results but with smaller numbers: for example, 50:24 crossing intermediate boundaries of 1.5 and 3.0 km around major oil depots. These effects were evident for both major and minor oil storage facilities. An asymmetry displayed by fossil fuelled power stations at 8.0 km was nearly matched by an equivalent random PC object set, but a shorter range intermediate boundary again showed significant differences.

Table 3   Short range asymmetries among birth-death address pairs* in the vicinities of non-VOC sources

Hazard type	Sites	Births:x:deaths†	Birth-nearest: death-nearest	p Value (within pair)
Iron/steel casting	18	52:2:31	97:49 (1.98)	<0.001
Gasworks	98	348:3:296	345:266 (1.30)	<0.005
Soap manufacture	21	46:2:20	89:60 (1.48)	<0.025
Brewing	33	110:3:78	126:85 (1.48)	<0.01
Detergents	33	81:2:36	162:95 (1.71)	<0.001
Harbours	80	145:3:101	161:102 (1.58)	<0.001
Potteries	40	50:2:28	93:60 (1.55)	<0.01
Railyards	105	262:3:221	284:225 (1.26)	<0.025
Zinc casting	43	71:2:46	149:98 (1.52)	<0.005
Hospitals	293	426:2:339	634:463 (1.37)	<0.001

*Birth and death 1.0–6.0 km apart : both within 10.0 km of a hazard. †"x" is the intermediate boundary in km. 160:3:128 means that 160 births and 128 deaths were within 3 km of a hazard site.

Table 4   Outward/inward migration asymmetries near combustion sources*. Crossings of neutral zone between two intermediate boundaries

Hazard type	Sites	Neutral zone (km)	Birth-nearest: death-nearest	P1†	P2†
Combustion sources‡	406	1.5 to 3.0	345:192 (1.80)	<0.001	<0.01
Hospitals (E and W)	293	1.5 to 3.0	338:201 (1.68)	<0.001	<0.05
Control PCs	350	1.5 to 3.0	266:209 (1.27)	<0.02	—

*Birth and death both within 20.0 km of a hazard. Migrants are included if one address is inside the neutral zone inner boundary and the other is outside the outer boundary. †P1 refers to primary asymmetry; P2 compares it with the random-PC ratio. ‡Gasworks, iron/steel foundries, steelworks, zinc casting, aluminium casting, railyards, docks/harbours, potteries, cement works, power stations, crematoria.

Table 6 gives other outcomes with greater difficulties of discrimination. The migration asymmetries were statistically significant but those at greater distances were matched or nearly matched by corresponding random PC object sets; while the smaller samples at closer ranges were statistically marginal. The raised risks around plastics works, car battery manufacturers, aluminium casting works, and in the near vicinities of airfields can probably be regarded as confirmed.

LINEAR HAZARDS
Previous studies showed raised cancer densities within 4 km of railways and of motorways outside London. (London rail networks were too complex and the M25 London Orbital Motorway was not then in use). Migrations near these same linear systems were now studied among children who were born and who died outside the London "box" (500–560E;160–200N).

Table 7 confirms the earlier railway results. It shows numbers of births and deaths at different perpendicular rail distances. Both the total samples of migrants and the subsets selected by various zonal contrast criteria, showed asymmetries. They were stronger at shorter ranges. This table displays results for the full set of migration addresses, but similar results were also obtained for several restricted sets. In clear contrast with the railways, and also with our earlier findings, the motorways showed no migration effects. Many different acceptance criteria were tried but no asymmetries were found.

TUMOUR TYPE SPECIFICITIES
The above presentations relate to "all cancers", but all analyses were disaggregated into the 11 cancer and leukaemia sub-types (above) and to two major subgroupings ("RE", the first five in the list: and "Solid"). Many sub-tabulations suggested specific exposure cancer relations but the numbers were usually small and their reality hard to assess. The issue was therefore tested more formally, although indirectly, by testing migrants with each specific cancer against birthplace objectsets derived from children with the same and with different tumours. The object sets were derived jointly from non-migrant and from non-matching migrant children.

These studies included migrants with less certain addresses but restricted the birthplace object sets to "certain" coordinates. This was to expunge the effects of inadvertent local groupings of poorly resolved coordinates. These studies showed clear differences between the Solid and the RE cancers. The migrating Solid cancers reacted more strongly with both the Solid and RE object sets than did the RE migrants; and the Solid object set elicited a stronger reaction from both migrant groups than did the RE object set. The differences were significant and consistent for different migration distances, as well as for different age groups.

For example, the Solid object set elicited outward/inward migration ratios of 1236:714 (1.731) overall: divided between Solid and RE

Table 5 Outward/inward migration asymmetries near oil and power plants. Crossings of neutral zone between two intermediate boundaries

Hazard type	Sites	Neutral zone radii in km*	Birth-nearest: death-nearest	P1†	P2†
Oil refining‡	25	25	769:640 (1.20)	<0.001	—
Benzene refining	6	10	69:71 —	NS	—
Oil depots§	77	12	1221:841 (1.45)	<0.001	NS
		1.5 to 3.0	219:77 (2.84)	<0.001	<0.001
		1.0 to 2.0	176:33 (5.88)	<0.001	<0.001
Random-PCs	77	12	1267:893 (1.42)	<0.001	—
		1.5 to 3.0	94:92 (1.02)	NS	—
		1.0 to 2.0	58:62 (0.94)	NS	—
Oil terminals¶	116	6	865:713 (1.20)	<0.001	NS
		1.25 to 2.5	114:65 (1.75)	<0.001	<0.01
Random PCs	116	6	1072:796 (1.35)	<0.001	—
		1.25 to 2.5	117:115 (1.02)	NS	—
All oil‡§¶	218	6	1380:1166 (1.18)	<0.001	—
		1.5 to 3.0	349:167 (2.09)	<0.001	<0.001
		1.0 to 2.0	242:82 (2.95)	<0.001	<0.001
Random PCs	218	6	1563:1070 (1.46)	<0.001	—
		1.5 to 3.0	342:309 (1.11)	NS	—
		1.0 to 2.0	210:209 (1.00)	NS	—
Power stations	69	8	1113:722 (1.54)	<0.001	NS
		1.5 to 3.0	99:49 (2.02)	<0.001	<0.025
Random PCs	69	8	963:667 (1.44)	<0.001	—
		1.5 to 3.0	91:89 (1.02)	NS	—

*Migrants are included if one address is inside inner boundary of the neutral zone and the other is outside the outer boundary. Some instances (single integers) use a single intermediate boundary, without a neutral zone. Outer limits vary: 20 or 30 km. †P1 refers to primary asymmetry; P2 compares it with the random-PC ratio. ‡Refineries plus major petrochemical works. §Major named oil company storage sites/terminals. ¶Railside and factory storage tanks.

migrants as 664:351 (1.892) and 572:363 (1.576). The RE object set gave 1228:764 (1.607) overall: divided between Solid and RE migrants as 644:352 (1.830) and 584:412 (1.417). These contrasts are for transfers between an inner zone of 1.0 km radius and an outer zone between 2.0 km and 10.0 km. The narrow inner zone effectively separates these results from density related interactions. Other boundaries gave similar results. "Certain address" migrants, tested in the same way, gave a similar pattern.

At the level of the individual tumour types, however, the results were non-specific. None of

Table 6 Outward/inward migration asymmetries near other point hazards. Crossings of neutral zone between two intermediate boundaries

Hazard type	Sites	Neutral zone radii in km*	Birth-nearest: death-nearest	P1†	P2†
Car batteries	14	5.0 to 10.0	186:105 (1.77)	<0.001	<0.001
Random PCs	14	5.0 to 10.0	69:63 (1.19)	NS	—
Crematoria	64	7	1238:795 (1.56)	<0.001	NS
		2.5 to 5.0	316:225 (1.41)	<0.001	NS
Random PCs	64	7	890:606 (1.47)	<0.001	—
		2.5 to 5.0	179:138 (1.30)	<0.025	—
Aluminium casting	24	1.5 to 2.0	76:47 (1.62)	<0.025	
Steelworks	44	12	542:438 (1.24)	<0.005	NS
Random PCs	44	12	1080:727 (1.48)	<0.001	—
Cement works	12	45	964:679 (1.42)	<0.001	NS
Random PCs	12	45	934:720 (1.30)	<0.001	—
Plastics works	23	3.0 to 5.0	76:43 (1.77)	<0.005	
Airfields	190	2.0 to 10.0	32:17 (1.88)	<0.05	
Tyre makers	23	7	200:148 (1.35)	<0.01	

*Migrants are included if one address is within the neutral zone inner boundary and the other is outside its outer boundary. Some instances (single integer formats) use a single intermediate boundary. †P1 refers to primary asymmetry; P2 compares it with the random-PC ratio.

*Table 7 Migrations near linear hazards: road and rail**

Hazard type	Sites†	Neutral zone radii in km‡	Birth-nearest: death-nearest	p Value
Railways	22590	inside 20	4150:3170 (1.31)	<0.001
		1.0 to 2.0	1681:771 (2.18)	<0.001
		0.25 to 2.0	597:249 (2.40)	<0.001
Motorways	2766	inside 20	2872:2795 (1.03)	NS
		4	964:973 (0.99)	NS
		1.0 to 3.0	229:240 (0.95)	NS

*Excludes cases born or died in the "London-box" (500–560E;160–200N). †Number of the digitised points in the network. ‡"Inside 20" indicates acceptance of all couplets within 20 km. Integer format indicates use of a single intermediate boundary.

the 11 type specific object sets elicited a greater than average reaction from its corresponding type specific migration group; and the greatest outward:inward ratios for type specific migration groups were not specifically associated with their corresponding object sets. There were no obvious subgroup associations apart from the two (RE and Solid) already noted.

Tumour specific correspondences of geographically adjacent cases would have indicated local discharges of tumour specific toxins, but only indirectly, and without showing what they were. Furthermore, their apparent absence might reflect only the inefficiencies of the indirect approach. Direct type specific searches against particular exposures were therefore conducted using the same zone separations as above (≤1.0 to >2.0 km), but without the 10 km limit. The "542" control object group shown in table 2 was used for comparison and several strong associations were detected with the main cancer subgroups. Most of them reinforced the general dominance of the Solid tumour migration ratios; although the VOC sources, hospitals and gasworks had their greatest effects upon the RE group. Nowhere, however, was there direct evidence of individual one to one type specific toxicities.

The most striking source was the group of 77 large oil storage facilities cited in table 5. They gave an overall outward/inward ratio of 5.33; 5.88 (100:17) for the migrating Solid cancers and 4.75 (76:16) for RE. Individual cancers showed more extreme ratios, but with small numbers. They included 14:1 for fatal benign tumours, 13:1 for neuroblastoma, 8:1 for nephroblastoma, and 33:6 for CNS tumours.

Gasworks and "all oil sources" elicited notable migration ratios for "Other Solid" tumours (29:6 and 20:3). Most were soft tissue sarcomas. A small group of 22 manufacturers of halogenated hydrocarbons showed a strong apparent association with CNS cancers. The outward/inward ratio across a radius of 5.0 km was 28:9, contrasting with 111:89 for all other cancers. There were other extreme ratios as well; but none of them can be treated as reliable signs of tumour specific toxicities.

Discussion

Our two purposes were to develop a new method of investigation and to retest the prior hypothesis that childhood cancers can be initiated early in life by pollutants originating from industrial sites. Both objectives sprang from earlier demonstrations that these diseases occurred in geographical clusters: that there were high case densities around suspect toxic sources relative to numbers of PCs: and that among children who moved house the proximity patterns were stronger for their birthplaces than for their deathplaces. The new method is based upon migrations among children who later develop cancer and it depends upon the premise that the risks from nearby early exposures differ from later ones. Population numbers are not used and it is free of uncertainties arising through inexact estimations of numbers of children close to suspect sources. Of two

methods tested, a case centred technique relying upon the closest approaches to particular hazard types proved the more powerful and flexible. In particular, different patterns of risk diffusion could be tested against subsets of children who had migrated in the most informative and context sensitive ways. This approach should prove useful in many other situations.

The main caveat is the possibility that the birth-death asymmetries might reflect general migrations in the whole child population and not just those with cancer; especially as observed cancer migrations in the vicinities of randomly selected PCs imply a movement from more PC dense to less PC dense areas. In the early years much substandard and war damaged housing was replaced, often involving city centre area demolitions with suburban and out of town reconstruction. These changes involved many young families and must have accounted for some outward migrations.

However, the cancer migration asymmetries around cancer birthplaces and around many suspect sources, greatly exceeded those around random PC object sets and around non-polluting industries. They also continued into the later years and were most striking at close ranges, strongly indicating the presence of local cancer toxic sources. A sufficient demographic explanation would have to invoke processes much more specific than a general rehousing programme.

A more plausible demographic model might propose a selective migration of young families to less dense environments, balanced by a reverse movement by others to their vacated addresses. However, the numbers of cancers moving out of town centres were not large; the 10 km radius circles around 48 cathedrals covered the centres of many major cities and a major slice of the national population, yet detected a net outflow of only 226 (1%) cancer cases. In addition, the increasing ratios within increasing circles around the random residential PCs, suggest a progressive incorporation of focal hazards around sites that are not intrinsically toxic. This is the inverse of the range associated decreases round cancer birth addresses and round many nominated hazards: and readily interpreted as a progressive blurring of an intrinsic focal effect as less toxic locations are incorporated.

The migration-based results were generally similar to our earlier case density-based findings. Most discrepancies of detail probably reflect the improved efficiency of the new method. Positive migration findings that contradicted earlier negative ones (electroplating, tyre making, gasworks, soap making) can probably be accepted. In the reverse direction, the only major unexplained contradiction between the two measurements related to motorways. Migration patterns did not support the earlier positive results. Industrial confounding offers a possible explanation; other hazards, located some distance apart, might extend the range of a local case density excess, yet compete with an inverse directional effect. The circum airfield

findings (table 6) seem similar. Both associations remain uncertain.

There were several indications that the proposed cancer initiating events must be limited to the prenatal period or to the very earliest part of postnatal life. Firstly, migration effects did not vary with age at death, and the shorter birthplace sojourns—identifiable with the younger migrant deaths—were no less effective than longer ones. Likewise, birthplace object sets derived from migrant and non-migrant births (that is, short and long stay cases) elicited similar outward/inward ratios among migrants. Finally a supplementary study showed that mean distances to nearest hazards, or to nearest random PCs, were no different in migrants dying earlier or later. If an extended postnatal cumulative exposure were necessary, we should have expected a more intense exposure and therefore a shorter hazard distance, among those exposed for a shorter time.

The question of the tumour specificities of different exposures was not resolved. Overall, the Solid tumours gave more extreme migration ratios, and were associated with more hazardous birthplaces, than the RE group, but each type of cancer birthplace elicited migration asymmetries among all cancer types. Hospitals, gasworks, and VOC sources elicited greater migration ratios among RE cancers, while combustion sources and oil storage installations preferentially affected the Solid tumours; but each showed strong cross effects. There were several very powerful Solid tumour reactions in the near vicinities of major oil storage facilities and gasworks; and halogenated hydrocarbons showed a particular link with CNS tumours. However, the reducing numbers resulting from specific short range searches made it difficult to distinguish real effects from sampling variations.

This paper confirms our earlier conclusions. Multiple focal toxic sources are probably responsible for many birth time or pre-birth cancer initiations. The effect is probably mediated through atmospheric pollutants of two main kinds, namely (a) various volatile organic substances, and (b) products of combustion. If this is so then there must also be many hazardous industrial sources beyond those tested; also many smaller local sources, each less productive than a major industrial site and effective only at short ranges. These latter may account for the fine scale geographical clustering demonstrated in earlier studies and by migration patterns associated with high local population densities. Parental occupational exposures may play a greater part here than in relation to migrations observed around large and sometimes sparsely manned industrial sites.[4] Such small density related risk zones may be located in semi-urban and in rural areas as well as in large towns; perhaps reconciling apparently conflicting observations that childhood leukaemias occur preferentially in rural zones[6–8]; yet within high density populations.[5] Clinical presentation is probably provoked by factors other than those responsible for initiation. In the case of the leukaemias these factors are probably

infective, as evidenced by the close spacetime patterns of some onset clusters.

The proposed chemical nature of the initiating exposures, the age of greatest sensitivity (corresponding with the embryonic tumour cell types), and the age distribution of onsets (analogous with those following prenatal radiation inductions), are consistent with the known epidemiological and animal experimental facts of cancer causation. The main problem of interpretation relates to the very low observed concentrations of carcinogenic substances detected in the atmosphere. These findings seem to demand substances of unprecedented toxicity, or some means through which low atmospheric exposures might be increased. In fact, both are well attested. For the first, the chemical family of the polychlorinated biphenyls (PCBs) and dioxins have been widely cited. For the second, the mother may serve as a cumulative filter, and initiation in the fetus would then depend upon prolonged prior maternal exposure to persistent toxic materials. Atmospheric dispersion could reach the mother directly or through food chain contamination. It has been shown repeatedly (for example, Lackmann, *et al*[9]) that the greater part of the childhood body burden of PCBs is acquired transplacentally or through breast milk. Our findings suggest that many cancer initiations in children may indeed depend upon these mechanisms and upon these or other volatile substances.

Appendix 1
SAMPLING, IDENTIFICATION, AND LOCATION OF HAZARDS

The existence and locations of the environmental hazards were determined through several routes. Many, such as oil refineries, gasworks, airfields, power stations, steelworks, small country towns and crematoria were identified and located through direct searches of Ordnance Survey (OS) maps and street atlases, on which their natures were directly declared. Others such as motorcar works, nuclear installations, benzene refineries, oil storage sites, hospitals, cathedrals, football grounds, and TV transmitters were first identified in specialised lists and charts, and OS map references then taken directly. Smaller factories were first identified in classified business directories,[10 11] and then located either on OS maps or through translation of their postcodes to map references via the Central PC Directory. Where recently allocated "business" PCs could not be traced in our "residential" directory we used the alphabetically nearest residential PC (last character) as the index position. Sites that could not be located adequately were excluded. "Company House" business addresses unattached to

factories were also excluded. Some locations were reconfirmed or amended through site visits. Detected errors, for example where the exact site on an industrial estate could not be discriminated on the map, were typically between 100 and 300 metres.

Except for the hospitals the major sites detected through map searches and specialised lists were essentially complete although a few gasworks and crematoria (for example) might have been missed. Hospitals were limited to England and Wales and comprised only the main administrative headquarters of hospital groups. No attempt was made to ascertain the remainder of the 2400 or so hospitals in the whole country. From a general knowledge of their names and locations, confirmed by their architectural structures as shown on OS maps, we could conclude that almost all of them were old hospitals or else situated on old hospital sites, and had been there throughout the whole of the birth years. Lists of particular hazard types taken from business directories often consisted of the full available set, typically 10–50 addresses, while others comprised the first 40–50 addresses (excluding the non-locatable) taken in alphabetic order. For electroplating works, a supplementary alphabetic sample was taken to test an indeterminate result.

Linear cartographic features—railways and motorways—were digitised from OS motoring atlases using a digitising tablet. Their accuracy was checked against larger scale maps. The tablet registered strings of point coordinates, and labelled 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 initially digitised in 11 separate regions: with London excluded because of its density and complexity. For the present report, these regions were consolidated. Disused lines were not examined. Motorways outside the recently completed M25 London Orbital (which was itself omitted) were digitised as a single set. High tension power lines appropriate to the times were not available on suitably scaled maps and were not examined.

Some sites (including all power stations, steel works, refineries, airfields, municipal gasworks, railways and almost all the hospitals and car factories) had evidently existed throughout the period of the survey and back as far as the earliest births. Many were no longer in use. Others such as the major TV transmitters, nuclear power stations and motorways were more recent and had existed for only for the later period. This may also apply to particular sources among the mail order centres, powder coating factories, solvent or pesticide or tyre manufacture, breweries, and furniture making and food preparation factories. Some newer factories may have been built on the sites of old ones, but not necessarily of the same kind. The maps used for identifying and locating hazards (OS Landranger 1:50 000) were issued in the 1980s and carry no note of the antiquity of the marked sites but comparisons with older 1 inch:mile series (early 1960s), where they were

available, showed few substantive changes, and these chiefly in relation to roads. Sites in new towns were not used; nor schools, nor coalmines. At the extremes of transience sites of municipal incinerators and petrol stations were apparently too mobile even to deserve a map record. In the absence of readily available early records, our intentions to use these sites were abandoned.

Appendix 2

ALTERNATIVE SEARCH STRATEGIES

Two modes of analysis (1 and 2) were investigated. Mode-1 was an extension of the method used in our previous paper. A list of hazard coordinates (the "object set") was scanned; for each hazard site a secondary scan of migrating children identified those whose two addresses (or optionally, either address) were within a stated radius of a source (although not necessarily the same source). The birth and death addresses were distributed to successive peri-hazard annuli and the two distributions compared. Cases failing to meet the acceptance criteria were excluded, and some address pairs were included more than once in relation to several sources. Mode-1 can be described as "hazard centred".

In mode-2 the address pairs of the migrants were identified first and the object set then searched for sources within a prescribed distance of either address. (Or, optionally, both addresses). This sequence is therefore "case centred". Here, however, statistical examinations were limited to the shortest identified hazard distances at birth and at death. These two shortest distances might refer to different sources; or, optionally, might be constrained to the same source—the source of closest approach to either address. Cases failing to meet the acceptance criteria were excluded but none was entered more than once. Particular sources might be excluded or might be used one or more times. The addresses were treated as linked pairs and examinations were restricted, as appropriate, to subsets whose migration dis-

tances (or absolute radial components) were between set limits; or whose addresses straddled intermediate limits (see text).

Under both modes, and whatever the options or the selective entry conditions, both addresses were invariably used; and the null hypothesis was always one of symmetry between the death to hazard and birth to hazard distances. For both search stategies, the separate distributions of the two hazard distances were compiled and compared; but the case centred approach also supplied a distribution of birth-death intra-pair differences between the two shortest values: that is, a positive or negative radial component of migration. It enumerated those cases where the shortest birth distance was less than the shortest death distance: and the opposite; and calculated the statistical significance of the difference. It calculated the mean radial migration distance (+/-), and the standard error of its separation from zero. Finally, the case centred approach represented migratory movements as vectors drawn on circular maps, centred on a coincident registration of the sources of closest approach.

1 Knox EG, Gilman EA. Leukaemia clusters in Great Britain. 2. Geographical concentrations. *J Epidemiol Community Health* 1992;**46**:573–6.
2 Knox EG. Leukaemia clusters in childhood: geographical analysis in Britain. *J Epidemiol Community Health* 1994;**48**:269–76.
3 Knox EG, Gilman EA. Spatial clustering of childhood cancers in Great Britain. *J Epidemiol Community Health* 1996;**50**:313–19.
4 Knox EG, Gilman EA. Hazard proximities of childhood Cancers. *J Epidemiol Community Health* 1996;**51**:151–9.
5 Muirhead CR. Childhood leukaemia in metropolitan regions of the United States: a possible relation to population density? *Cancer Causes Control* 1995;**6**:383–8.
6 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**:577–92.
7 Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. *Br J Cancer* 1995;**71**:1–5.
8 Kinlen LJ, John SM. Wartime evacuation and mortality from childhood leukaemia in England and Wales in 1945–9. *BMJ* 1994;**309**:1197–202.
9 Lackmann G-M, Goen T, Tollner U, *et al*. PCBs and HCB in serum of full-term German neonates. Lancet 1996;**348**:1035.
10 *Kompass: CBI-UK*. East Grinstead: Reed Information Services, 1993.
11 *Kelly's*. East Grinstead: Reed Information Services, 1994.