SOMERSET HEALTH AUTHORITY

LEUKAEMIA INCIDENCE IN SOMERSET

WITH PARTICULAR REFERENCE TO HINKLEY POINT

BY P D EWINGS PhD STATISTICIAN

C BOWIE MRCP MFCM DIRECTOR

DEPARTMENT OF COMMUNITY MEDICINE

DECEMBER 1988

CONTENTS

INTRODUCTION

1. Background
2. National Perspective
3. Local Perspective
4. Radiation Risks

ANALYSIS I - LEUKAEMIA DISTRIBUTION IN WEST SOMERSET 1971-1987

1. Introduction
2. Number of Cases
3. Trends over Time
4. Incidence in the West of Somerset compared to LRF rates
5. Distribution in Electoral Wards
6. Childhood Leukaemias

Maps

ANALYSIS II - LEUKAEMIA REGISTRATIONS IN YOUNG PEOPLE IN THE VICINITY OF HINKLEY POINT, 1959-1986

1. Introduction
2. Case Ascertainment
3. Validity of the Data
4. Age Range Studied
5. Time Periods Studied
6. Geographical Boundaries Selected
7. Actual Areas Included
8. Population Estimates
9. Types of Cancer Studied
10. Statistical Methods
11. Registered Cases of Leukaemia and NHL
12. Leukaemia and NHL Incidence - main result
13. Sensitivity to Choice of Geographical Boundary
14. Time Trends

Maps

Appendix

DISCUSSION

1. Why Two Analyses ?
2. Summary of Results
3. Limitations of Analyses
4. Possible Explanations for Results
5. Conclusion

ACKNOWLEDGEMENTS

REFERENCES

SUMMARY

At its meeting of 30 November 1984, the Health Authority received a report on leukaemia in the west of Somerset. A second report was presented at the 30 January 1987 meeting. This third report contains the results of two further analyses examining leukaemia incidence in Somerset.

Analysis I

Analysis I updates the results presented in the two previous reports. The geographical distribution of leukaemia cases throughout the catchment area of Musgrove Park Hospital (MPH) is examined for the study period 1971-1987. Particular attention is focussed on incidence in electoral wards around Hinkley Point nuclear power station, but no specific hypothesis relating to the station has been formulated or tested.

Analysis I confirms the findings of the two previous studies. There is a high rate of leukaemia incidence (all ages) in the MPH catchment area, compared with rates found by the Leukaemia Research Fund (LRF) for a large part of the country. LRF rates are only available for the three-year period 1984-1986. For the 17-year period 1971-1987, the local rates are 24% higher than LRF rates for 1984-1986. Limiting the analysis to the three-year period for which LRF rates are available reveals a local rate for 1984-1986 that is 61% higher than the LRF rate.

The cases of leukaemia were spread across the area studied, with no evidence of "clustering". The electoral ward in which Hinkley Point is situated does have a high rate for those leukaemias thought to be related to radiation, but such a rate is found in other wards in the study area, and is, therefore, not exceptional in this context. The cases of childhood leukaemia were too few to enable a full analysis to be undertaken. There appears to be no evidence of clustering, with cases scattered throughout the MPH catchment area. The ward containing Hinkley Point has the highest rate of all 91 wards in the study area for the under 15 age group, but this rate is based on just two cases and no conclusions can be drawn from such a small number.

Analysis II

Analysis II sets out specifically to test the hypothesis that there is an increased incidence of leukaemia among young people in the vicinity of Hinkley Point. The methods used for this analysis follow as closely as possible those employed by the Committee on Medical Aspects of Radiation in the Environment (COMARE) in its study of leukaemia incidence around the Dounreay nuclear installation in Scotland. This ensures a degree of objectivity in that the methodology has been established in advance. Analysis II uses data from the national cancer registry system for the study period 1959-1986. Particular attention is focussed on leukaemia incidence since 1964, when Hinkley Point "A" station commenced operations.

Analysis II identifies high rates of leukaemia and non-Hodgkin's lymphoma (NHL) among under-25 year olds in the Somerset Health District (SHD).

However, higher rates still are found in the vicinity of Hinkley Point. The main findings are:-

1. For 1964-1986, a specific area around Hinkley Point had 19 registered cases, 9 more than could be expected (ie 10) from national rates. A formal statistical test shows that the high rate for the Hinkley Point area is unlikely to have arisen simply "by chance".

2. For 1964-1986, the remainder of SHD had 118 registered cases. This is 18 more than could be expected from national rates, and is fairly unlikely to have arisen by chance.

3. If the rate in the Hinkley Point area had been similar to the high rate in the rest of SHD during 1964-1986, 12 cases (not 10) would have been expected. Statistically, the seven extra cases could have arisen by chance, with the rate for the Hinkley Point area just reflecting the high rate throughout SHD.

4. Analysis of separate five-year time periods reveals the following:-

   (a) For the five-year period 1959-1963 before the "A" station was commissioned, the rate throughout SHD (including the Hinkley Point area) was higher than the national rate.

   (b) For the ten-year period 1964-1973 following the commissioning of the "A" station, there was a markedly high incidence in the Hinkley Point area while the rate for the rest of SHD was close to the national average. In particular, there were nine cases registered for the Hinkley Point area in the five-year period 1969-1973, when only two could be expected from national rates. This difference is unlikely to have arisen by chance.

   (c) Subsequent to 1973, rates for the Hinkley Point area have generally been slightly lower than national rates, while those for the rest of SHD have been slightly higher than national rates.

Analysis II thus identifies an unusual pattern of leukaemia incidence in the vicinity of Hinkley Point. A relatively high rate exists for the period 1964-1986 but the excess cases are concentrated in the period 1964-1973; after 1973, the rate is unremarkable. There is no ready explanation for this pattern. If radioactive emissions from Hinkley Point are responsible, large unreported releases would need to have occurred in the 1960's. This possibility needs to be explored. There are other possible explanations but current knowledge about the causes of leukaemia is insufficient to offer definite answers.

INTRODUCTION

1. Background

Two reports have previously been presented to the Health Authority examining leukaemia incidence in the west of Somerset - one in November 1984[1] and the other in January 1987[2]. The reports concluded that although there appeared to be a high incidence of leukaemia in this area, there was no evidence to suggest "clustering" of cases in any particular location, and no known reason for the relatively large numbers of leukaemias found overall. In particular, rates around Hinkley Point nuclear power station were not suggestive of an unusual incidence in that vicinity. In this third report, two further analyses examining leukaemia incidence in Somerset are presented.

2. National Perspective

Nationally, concern continues to be voiced about the possibility of radioactive emissions from nuclear installations contributing to a raised incidence of cancer (and in particular, leukaemia) in their vicinity. Such is the complexity of the issues involved that it is still disputed whether unduly high incidences <u>do</u> occur near nuclear plants. Claims and counter-claims abound in the literature, and it is difficult for the reader to establish the accuracy, relevance and importance of individual reports, other than to accept the interpretation of the findings presented by the authors themselves. This is particularly difficult when, as often happens, different authors draw different conclusions from the same data.

Clear interpretation is often thwarted by considerations of a statistical nature. It is not difficult, for example, to find areas of very high leukaemia incidence if the geographical boundary, time period and/or age group is selected <u>after</u> the data have been collected and studied[3]. In the report by the independent advisory group chaired by Sir Douglas Black[4], the levels of leukaemia in the vicinity of the Sellafield nuclear establishment were investigated following a Yorkshire TV broadcast suggesting an unusually high incidence. The report confirmed that there was a "higher incidence of leukaemia in young people resident in the area", but expressed reservations about the selection of specific geographic areas : "if parishes are selected for study because cases of cancer are known to have occurred there, it is not surprising if the incidence of cancer in those parishes is found to be unusually high" (para 2.10). Indeed, as has been suggested elsewhere[5], the Sellafield data ought to be viewed as data that <u>generate</u> a hypothesis, rather than data for <u>testing</u> a hypothesis.

The Sellafield data have provided the foundation for a hypothesis that there may be an association between nuclear plants and cancer incidence; variations of this hypothesis have since been tested in relation to several installations and results reported accordingly[5-8].

The most recently published study is a report[9] by the Committee on Medical Aspects of Radiation in the Environment (COMARE), a body set up in 1985 "to assess and advise government on the health effects of ..... radiation ..." following a recommendation in the Black report. Its report examined the evidence for an increased incidence of leukaemia among young people living in the vicinity of the Dounreay nuclear establishment, following a report of high levels by the Information and Statistics Division of the Common Services Agency for the Scottish Health Services[8]. COMARE concluded that there was indeed "evidence of a raised incidence" and that taken in conjunction with the Sellafield data, the evidence tended "... to support the hypothesis that some feature of the nuclear plants .... leads to an increased risk of leukaemia in young people living in the vicinity of these plants". However, it is still not clear as to why there is an increased incidence. Reported levels of radioactive discharges suggest that these could not be responsible if conventional dose/risk estimates are used; however, these estimates of the relationship between dose and risk are somewhat unreliable, and refinements continue to be made[10-13].

The finding of raised leukaemia levels in the vicinity of the Sellafield and Dounreay nuclear plants is not necessarily evidence of a cause and effect relationship, although alternative explanations have been explored without success. The locations of nuclear power plants have much in common; they tend, for example, to be in rural situations, they are usually on the coast, there are common geological factors and the people in the surrounding area tend to have a relatively high social class distribution. Any one of these factors may, for some reason, be associated with a high incidence of leukaemia for quite different reasons than the presence of the nuclear plant. In such a case a spurious association would be generated between nuclear plants and high levels of leukaemia, with no cause and effect basis. Moreover, the fact that Dounreay and Sellafield are the only installations in the UK that reprocess spent fuel means that findings in those areas may not be comparable with findings in the areas around other nuclear plants. Clearly the process of gathering evidence needs to continue.

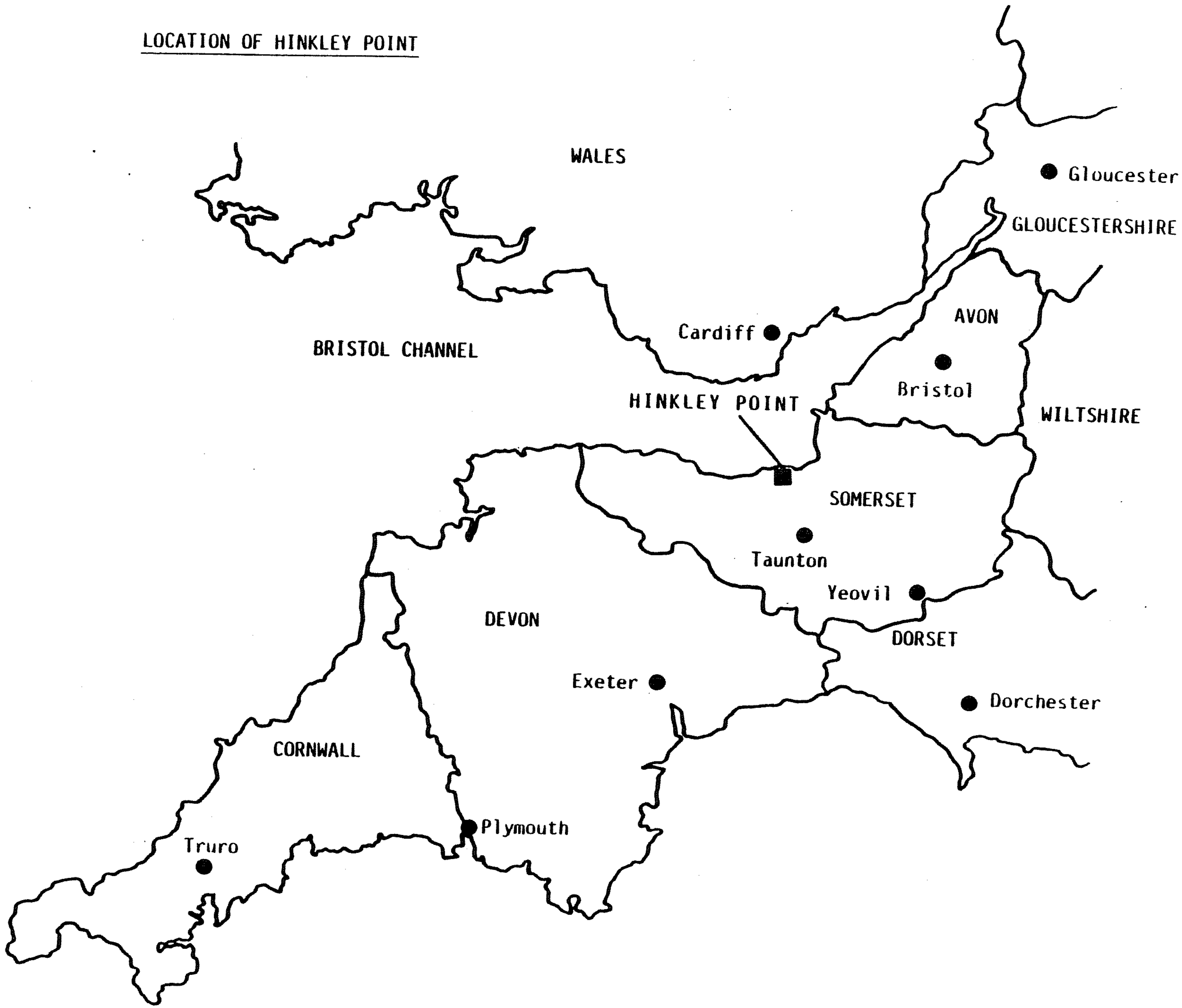
3. Local Perspective

Hinkley Point is a nuclear installation situated on the north coast of Somerset (see map). It contains two power stations, one containing two reactors of the older Magnox type (which went online in 1964) and one with two reactors of the AGR type (commissioned in 1977). The station is situated in a rural location with the nearest towns (excluding small villages) being Burnham-on-Sea (about 10 kilometres along the coast) and Bridgwater (about 13 kilometres inland).

The continuing public concern over the suggested links between nuclear establishments and increased levels of cancer has prompted the writing of this third report on leukaemia in the vicinity of Hinkley Point. Also, the public enquiry into the proposed building of a third station at Hinkley Point may wish to have access to information from the local health authority when examining possible health effects of radioactive discharges.

This report contains the results of two entirely different and separate investigations. The first is an update of the work presented in the second Somerset report[2], extended to include more recent data. The geographical distribution of leukaemia cases throughout the west of Somerset is examined generally, with particular reference to the location of Hinkley Point. However, no specific hypothesis relating to Hinkley Point is tested. The second analysis specifically tests the hypothesis that there is an increased level of leukaemia in young people in the vicinity of Hinkley Point, using the same methods as COMARE in its analysis of the Dounreay data.

LOCATION OF HINKLEY POINT



4. Radiation Risks

In the second of the two previous reports on leukaemia in the west of Somerset[2], an estimate was made of the Somerset population's exposure to radiation from all known sources. An estimate was then made of the number of leukaemias that might be expected to result from this level of exposure.

Estimated national levels of radiation exposure were used except where local information was available. For example, levels of terrestrial gamma radiation are available for a 10 km grid over the whole of the UK, enabling an estimate of exposure to be made for the Somerset population. Estimates of exposure to radiation from radioactive waste disposal were made using figures produced by the National Radiological Protection Board and relatively pessimistic assumptions about how much radiation the Somerset population is likely to be exposed to from this source.

As mentioned above, dose/risk estimates for radiation exposure are continually being revised. In the second Somerset report, a simple linear model of the relationship between dose and risk was used to estimate the number of leukaemias that might have arisen through radiation exposure. It was estimated that just 1.1 leukaemias a year in Somerset can be attributed to radiation under the assumptions used. One fifth of this figure is due to man-made radiation, but only a small fraction (0.011 per year, or one every 91 years) is attributable to the nuclear industry.

There is no reason to suspect that exposure to radiation in Somerset has changed dramatically since the second report was written. The recent revisions of the dose/risk relationship suggest that radiation might be responsible for two to three times more cases of leukaemia than previously thought, but the number attributable to the nuclear industry in Somerset's case, based on the assumed exposure levels, would still be very small. These reported levels are estimated to form less then one percent of the total radiation exposure for the Somerset population. Only if the assumed doses from Hinkley Point are drastically wrong (either by an underestimation of the emissions or by underestimation of the amount actually reaching people) could there be a significant contribution to total radiation exposure from this source. In other words, however many leukaemias might be attributable to radiation (using increasingly refined dose/risk models), only a small fraction are likely to be due to radiation from Hinkley Point, using current knowledge about exposure to radioactivity from the nuclear plant.

This third report does not re-examine radiation levels and likely numbers of leukaemias attributable to radiation from various sources. It is largely epidemiological in nature; it examines the numbers of leukaemias near Hinkley Point and assesses whether these numbers give cause for concern when compared to rates found elsewhere.

ANALYSIS I. LEUKAEMIA DISTRIBUTION IN WEST SOMERSET 1971-1987

1. Introduction

This analysis updates the results presented in a previous report[2]. In that report cases of leukaemia during 1971-1985 had been identified from two data sources - the South Western Regional Cancer Registry and the local haematology department at Musgrove Park Hospital (MPH), Taunton. The Cancer Registry data are not complete, especially for earlier years[14], and the local records provide details of additional cases. The use of both sources results in a more complete case ascertainment, but makes comparison with national cancer registry data untenable; a higher incidence of leukaemia found locally might be due solely to the enhanced method of data collection.

Instead, the number of leukaemias found in the MPH catchment (see map 1 at the end of this analysis) was compared with rates produced by the Leukaemia Research Fund (LRF). The LRF collects data from various sources (including cancer registries and local haematologists) over a large part of the country, as illustrated in map 2. Its data collection commenced in 1984, and provides a level of case ascertainment that should parallel that produced for the MPH catchment.

At the time of writing the second report, only 1984 rates were available from the LRF and these were used to assess leukaemia levels in the MPH catchment for the entire 1971-1985 period. Clearly, one year's data is not the best basis for comparing rates, as incidence can vary considerably from one year to the next. Moreover, it appears that for some types of leukaemia, incidence may have increased nationally since 1971. Thus the 1984 LRF rates may be an overestimate of actual national levels in the early 1970's, in which case their comparison with the early Somerset data should suggest relatively low rates for Somerset. In fact, high rates of leukaemia incidence were found throughout the MPH catchment, with an excess of cases far beyond the number that might be expected to occur by chance alone. It was concluded that, for some reason, there was an unduly high number of leukaemia cases reported in the west of Somerset. It was impossible to tell how much, if any, of the excess was due to assiduous reporting. There was, however, no evidence to suggest that cases of leukaemia were "clustering" together in any location; in particular, rates in electoral wards around Hinkley Point were generally no higher than elsewhere in the area studied.

The analysis presented in this section updates the previous report in two respects: Somerset data for 1986 and 1987 are included, and LRF rates for 1984-1986 are used. Using the three-year LRF rates will make the analysis less sensitive to the peculiarities that may be present for one year alone, but the problem of comparing "current" rates with historic local figures still remains. However, it is the distribution of cases within the area that is of particular interest. Following the previous report, high rates may expect to be found, but it is the geographical spread within Somerset that is the main focus of interest.

Through a continuing process of validation, errors discovered in the data used for the previous reports have been corrected; this has resulted in small discrepancies between the results presented in those reports and those given here.

2. Number of Cases

A total of 540 leukaemias has been established for the MPH catchment during 1971-1987. One particular type of leukaemia, namely, chronic lymphocytic leukaemia (CLL) is not thought to be related to radiation exposure, so it is useful to treat these cases separately in the analyses. In this report, the remaining leukaemias are simply referred to as non-CLL. Of the 540 leukaemias, 229 were the CLL type, leaving 311 non-CLL cases.

Table 1 gives an age breakdown of the cases. It can be seen that leukaemia is a disease that mainly affects the elderly. In the context of radiation exposure, attention is often focused on childhood leukaemia, and these will be examined separately. However, the numbers involved are so small that a proper statistical analysis is difficult to undertake.

Table 1 - Age distribution of 540 leukaemia cases in Musgrove Park Hospital catchment, 1971-1987

Age Group	Non-CLL	CLL	All Leukaemias
0 - 4	18	0	18
5 - 14	13	0	13
15 - 24	16	1	17
25 - 34	13	1	14
35 - 44	15	0	15
45 - 54	19	12	31
55 - 64	50	35	85
65 - 74	71	76	147
75 - 84	68	77	145
85+	28	27	55
Total	311	229	540

3. Trends over Time

In the previous report, time trends were examined by plotting the number of cases diagnosed in each year. This does not give a true reflection of incidence. The number of people in the MPH catchment has increased by over 10% since 1971 and the age structure has changed to an older population, so that more leukaemias can be expected to arise. Trends are therefore examined here with the size and age distribution of the population taken into account. This is undertaken using a common technique known as standardisation. Typically, this method uses age and sex-specific incidence rates in a "standard" population, and estimates of the local population size and structure, to derive "expected" numbers of cases locally. When looking at time trends, the first year (or some later year) is often taken as the standard, and expected numbers for other years are calculated by assuming the rates for the standard year continue to apply.

The ratio of the actual number of cases in a given year to the expected number so calculated is known as the standardised incidence ratio (SIR); it is usually expressed as a percentage by multiplying by 100.

As a result of the relatively small number of cases occurring in any one year, rates for a five-year period centred around the 1981 census (ie 1979-1983) are used here as the standard. The rates are calculated by dividing the average annual number of cases in each age/sex group over the five-year period by the 1981 census population figure for that age/sex group. Expected numbers and SIR's for all 17 years are then calculated accordingly, with the 1979-1983 period as a whole having a baseline index SIR of 100.

Figures 1 and 2 show the trends in the SIR's over the period studied for non-CLL and all leukaemias respectively. Figures 3 and 4 show the trends for leukaemias in young people; the 0-14 age range is conventional for childhood cancer analysis but 0-24 is also often used if the number of cases is small. Figures 1 and 2 show that there has been a gradual increase in reported cases of leukaemia in the MPH catchment since 1971. This observation applies to non-CLL cases as well as all leukaemias. Simple linear regression analyses of the data confirm the upward trends, with an average annual increase in the SIR over the period of 1.6 for non-CLL cases (b=1.6, 95% confidence interval=0.14 to 3.1) and 2.5 for all leukaemias (b=2.5, 95% confidence interval=0.80 to 4.2). It is impossible to say how much of this increase is due to better case ascertainment in later years. The large increase in all leukaemias in 1984 is thought to reflect the encouragement by LRF to report every single case. In previous years some cases may well have been missed, particularly CLL cases; CLL is generally a long-term illness that may often be discovered incidentally when a blood count is taken, and more cases have been identified in recent years through attempts to provide accurate assessment for the LRF study.

The trends in leukaemia rates in young people (figures 3 and 4) cannot be examined in the same way as those for all ages, because of the small numbers involved. There is a suggestion of a downward trend over time, but an appropriate test (based on Poisson errors) shows that this trend is not statistically significant.

4. Incidence in the West of Somerset compared to LRF rates

There are no national figures available that would allow a fair assessment of the level of leukaemia incidence in West Somerset compared to the country as a whole. The Cancer Registry system depends on voluntary notifications and is not complete[14,15]; the enhanced methods of data collection used for the MPH catchment have identified cases not on the registry and it would not be surprising if local rates so derived were higher than national cancer registry rates. Instead, as in the previous report, figures supplied by LRF are used for comparison.

The LRF study commenced in 1984, and data continue to be collected from a large part of the country (map 2). Information is obtained directly from haematology and other hospital departments throughout the study areas as well as from cancer registries. Case ascertainment levels should therefore be similar to that produced for the MPH catchment, at least since 1984; local case ascertainment before 1984 may be slightly inferior. Using 1981 census populations for electoral wards and rates calculated from a combination of Yorkshire, Trent and South Western regions (which LRF

TRENDS IN STANDARDISED INCIDENCE RATIOS FOR THE MPH CATCHMENT OVER THE PERIOD 1971-1987 (1979-1983 = 100)

FIGURE 1 NON-CLL LEUKAEMIAS

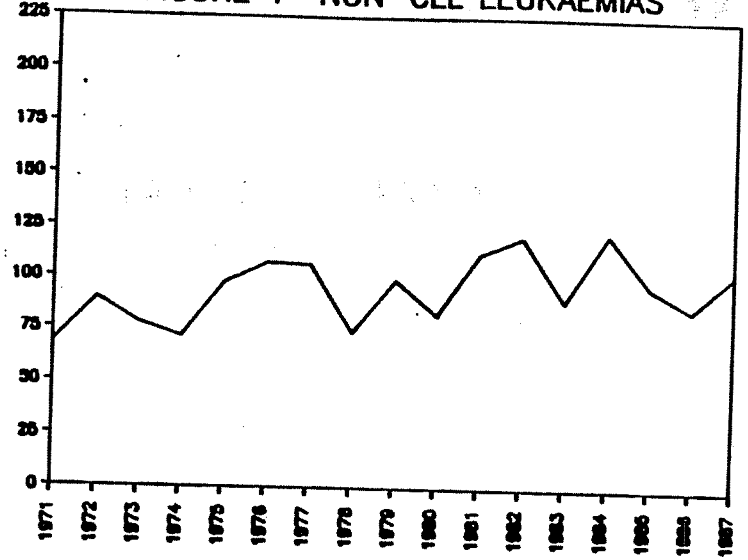


FIGURE 2 ALL LEUKAEMIAS – ALL AGES

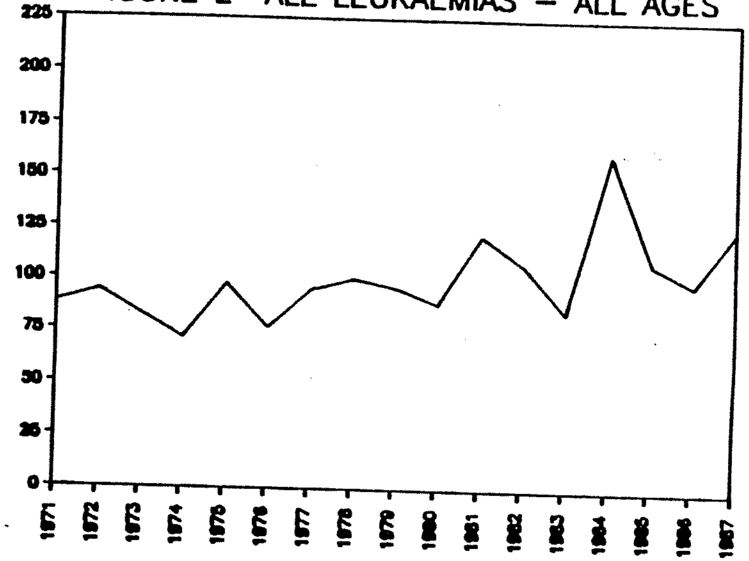


FIGURE 3 ALL LEUKAEMIAS – AGES 0–24

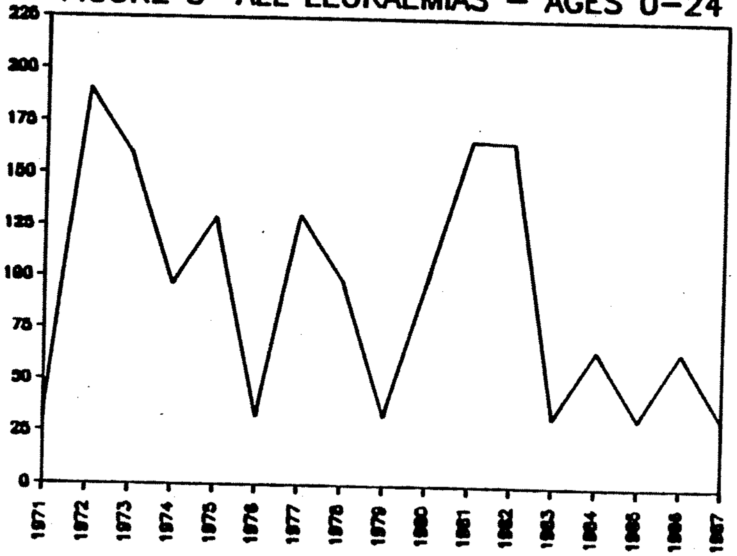
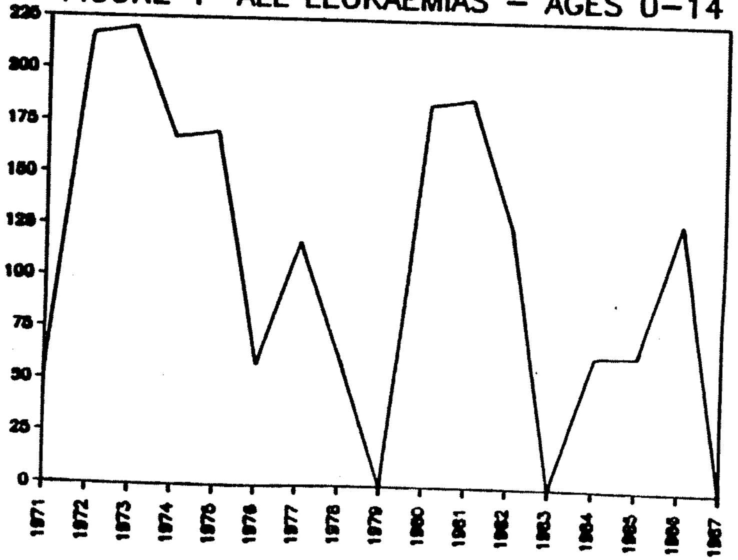


FIGURE 4 ALL LEUKAEMIAS – AGES 0–14



believe to be most reliable in reporting terms), LRF has supplied the expected numbers of leukaemias in the wards in the MPH catchment. These numbers are derived from the age/sex specific rates from the areas mentioned and the populations for the same age/sex groups in the wards of interest. Thus the age structure of the local population is taken into account and the actual number of leukaemias observed can be compared to the number expected.

The expected numbers supplied by LRF are based on the three year period 1984-1986. If incidence rates have remained fairly constant since 1971 then the average annual incidence in this three-year period can be used to estimate previous years' rates. In fact, as mentioned in the previous report, it seems that some types of leukaemia have increased in number since 1971. If this is so then the LRF 1984-1986 rates overestimate rates in previous years, and Somerset rates in early years should be low by comparison. Thus, comparisons of local figures for the entire period 1971-1987 with LRF figures for 1984-1986 will probably underestimate the relative incidence of leukaemia in the MPH catchment. It is not possible accurately to assess the national trends in leukaemia incidence (part of the increase is almost certainly due to better reporting), and therefore not possible to quantify the effect on the comparisons made here; for reference, local rates for 1984-1986 will be compared with LRF figures, in addition to the entire period under study.

Disease diagnosis is not an exact science and may become particularly unreliable when the patient concerned is very elderly. The distinction between certain types of lymphoma and leukaemia, for example, may depend on painful tests which might not alter treatment or prognosis for a very old patient; such tests are, therefore, not always carried out. Registration of a case from death certification only, may also be unreliable if, as is often the case for an old person, a post mortem is not performed. For this reason, the LRF omit data on persons aged 85 and over in calculation of expected numbers. Thus the figures presented in this analysis are for under 85's only.

As usual, observed and expected figures are compared here by calculating the standardised incidence ratio (SIR), which is the ratio of observed to expected numbers, multiplied by 100. The SIR for the MPH catchment is therefore a measure of deviation from the average of 100 for the LRF study areas. Even if local rates were similar to national rates it would be unsurprising to find a SIR different from 100, simply through natural variation. Statistical methods can be used to assess whether the deviation found could reasonably have arisen through random fluctuation (ie "by chance") or whether there appears to be a real difference. The relatively large numbers involved here mean that a standard approximation[16] can be used to construct a "95% confidence interval" (CI). This interval is a range of values that is likely (with 95% certainty) to encompass the "true" underlying SIR for local leukaemia incidence compared to LRF rates.

Table 2 makes the comparisons described above. The SIR of 124 for all leukaemias in the period 1971-1987 indicates that the 485 cases aged under 85 in the MPH catchment is 24% higher than expected from LRF rates. The 95% CI of 113 to 136 shows that this excess is unlikely to have arisen by chance and, as expected from the previous report, high rates in the west of Somerset are confirmed. For non-CLL cases, the excess of observed over expected is 30%.

Table 2 - Leukaemia cases in MPH catchment. Ages 0-84. Expecteds based on LRF rates 1984-1986

	OBSERVED (O)	EXPECTED (E)	SIR (100 X O/E)	95% CONFIDENCE INTERVAL, SIR
1971-1987				
non-CLL	283	217.5	130	115 - 146
CLL	202	173.3	117	101 - 134
All Leukaemias	485	390.8	124	113 - 136
1984-1986				
non-CLL	57	38.4	148	112 - 192
CLL	54	30.6	176	133 - 230
All Leukaemias	111	69.0	161	132 - 194

In view of the uncertainty involved in using 1984-1986 LRF rates to assess MPH catchment rates over 1971-1987, the local figures for just 1984-1986 are also shown in the table. The SIR's for this three-year period alone are higher than for the entire period, suggesting that the use of LRF rates for earlier years may have resulted in underestimation of SIR's for the full period. The result that leukaemia incidence in the west of Somerset is high is thus further emphasised. For 1984-1986 alone, where direct comparison with LRF rates is possible, there were 61% more leukaemias than expected in the MPH catchment. The excess for non-CLL cases is 48%. These are statistically significantly higher rates than for the LRF study areas as a whole.

5. Distribution in Electoral Wards

This section examines the geographical distribution of leukaemias in the MPH catchment by analysis of incidence in electoral wards. LRF has supplied expected numbers for the 91 wards in the catchment, and the cases described earlier have been assigned to ward according to address at time of diagnosis. As before, analysis is limited to persons aged under 85. In addition, two cases were unable to be coded to ward and have, therefore, been omitted, leaving a total of 483 cases.

For each ward, a SIR can be calculated as described above. Since the numbers involved for any one ward are small, the methods used earlier do not now provide an adequate approximation, and use is made of the more exact Poisson distribution. If a rare disease such as leukaemia is spread randomly throughout a population (geographically and temporally), then the number of cases occurring in a given area and time period behaves according to the Poisson law. This can be used to evaluate the extent to which an observed number of cases can be expected to deviate from the expected number purely through natural variation. An assessment can then be made about the likelihood that the deviation between observed and expected arose simply by chance.

In analyses of this kind, attention is usually focussed on those areas with a larger number of cases than expected. The "cumulative one-sided Poisson probability", P, is calculated, measuring the probability that such a large number of cases could have occurred by chance. The smaller the value of P, the more unlikely that the observation is a chance one. In this analysis the SIR and associated P value are calculated for each of the 91 wards in the MPH catchment. Since it is known from the previous section that incidence locally is high compared to LRF rates, it would not be surprising to find some wards with high SIR's, and indeed some small P values in any case can be expected by chance. What is of interest here is the geographical distribution of those wards with relatively high SIR's.

Tables 3 and 4 show those wards with the ten highest SIR values, and also other wards that are statistically significantly high at the 5% level (ie with a P value smaller than 5%). Because of the varying population sizes, it may happen that a small ward has a high SIR without being statistically significant while a relatively large ward may be significantly high with a smaller SIR. Table 3 relates to all leukaemias while table 4 deals with non-CLL cases only. The wards so identified are highlighted in maps 3 and 4.

Table 3 - Electoral Wards with Ten Highest SIR's and/or Statistically Significantly High SIR's. All leukaemias, 1971-1987. Ages 0-84.

WARD	1981 POPULATION	NO OF CASES OBSERVED	EXPECTED	SIR	P
Norton Fitzwarren	2015	7	2.41	291	.012
Holway	4413	23	8.04	286	.000012
Exmoor	711	3	1.23	244	.13
Glastonbury Street Benedicts	1389	6	2.48	242	.040
Creech St Michael	2279	8	3.54	226	.028
Galmington	5142	19	8.48	224	.0013
Parchey	2943	8	3.67	218	.034
Monument	1674	6	2.82	213	.067
Williton	2422	9	4.48	201	.039
Holnicote	692	3	1.53	196	.20
Street South	4922	14	7.26	193	.017

Table 4 - Electoral Wards with the Highest SIR's and/or Statistically Significantly High SIR's. Non-CLL Leukaemias 1971-1987. Ages 0-84.

WARD	1981 POPULATION	NO OF CASES OBSERVED	EXPECTED	SIR	P
Norton Fitzwarren	2015	6	1.47	407	.0041
Holnicote	692	3	0.80	373	.048
Parchey	2943	7	2.26	310	.0085
Exmoor	711	2	0.68	294	.15
Sowey	1182	3	1.06	283	.092
Quantock Vale	1894	5	1.84	272	.039
Pawlett & Puriton	2572	6	2.27	264	.028
Street South	4922	11	4.24	260	.0044
Galmington	5142	12	4.74	253	.0036
Monument	1674	4	1.59	252	.077
Curry Rivel	2529	6	2.46	243	.040
Williton	2422	6	2.47	243	.040
Ilminster	3682	8	3.63	221	.032
Holway	4413	9	4.42	204	.037

Map 3 shows that, for leukaemias as a whole, wards with relatively high SIR's do not form any obvious geographical pattern. In particular, there is no evidence of those wards around Hinkley Point having high rates compared to those elsewhere in the catchment. Map 4 again shows a wide distribution of wards with high SIR's when attention is focussed on non-CLL leukaemias. The ward containing Hinkley Point (Quantock Vale) is now in the top 10 SIR's and is statistically significantly high compared to LRF rates. However, table 4 shows that the rate for this ward is by no means exceptional in the MPH catchment.

6. Childhood Leukaemias - Distribution in Electoral Wards

In studies of this kind, particular attention is often focussed on leukaemias in young people. If radiation exposure is involved, the latent period between exposure and onset of disease is almost certainly shorter in children than in adults. Thus in a population exposed to increased levels of radioactivity, any effects are likely to be seen more quickly in children. Analyses of childhood cancer conventionally use the 0-14 age range, or perhaps the 0-24 range if numbers are small; both are used here.

The LRF has supplied expected figures for wards for all ages in total, so it is not possible to compare the number of childhood leukaemias in the MPH catchment as a whole with rates elsewhere. It is possible, however, to look at the distribution of these young cases throughout the catchment, by examining the number in each electoral ward. The numbers involved are too small to allow a full statistical analysis, but SIR's for each ward are calculated here using the entire MPH catchment as a standard. That is, based on the age-specific rate throughout the catchment, expected numbers (and hence SIR's) are calculated for each ward. The SIR for a ward therefore gives an indication of the ward's rate, compared to an average of 100 for the whole area under study.

There is a total of 48 cases aged under 25, of whom 31 are aged under 15. In the under 25 age group, the largest number of cases in any ward is three (one ward), while two cases was the maximum for the under 15 age group (seven wards). P values have been calculated, but they should be treated with extreme caution; a large standard population should normally be used when calculating SIR's and the use of the MPH catchment here as a standard, as well as the very small numbers involved, indicate that the P values should not be read too literally.

Those wards with the ten highest SIR's are shown in table 5 for the under 15's and table 6 for the under 25's. The distributions of all 31 cases aged 0-14 and all 48 cases aged 0-24 are shown by electoral ward in maps 5 and 6 respectively.

Table 5 - Electoral Wards with Ten Highest SIR's. Ages 0-14, 1971-1987. MPH Catchment used as standard.

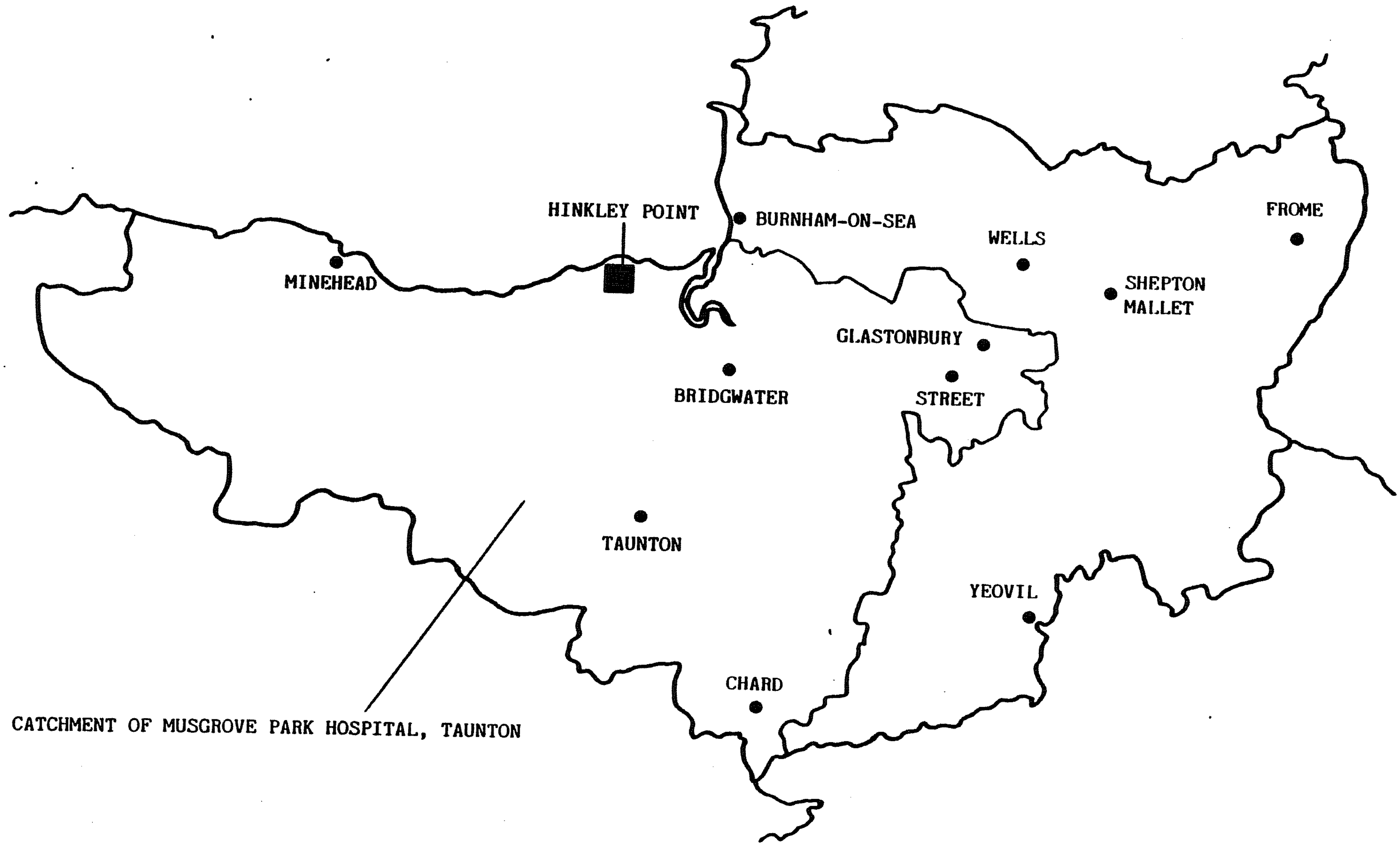
WARD	1981 POPULATION (0-14)	NO OF CASES	SIR	P
Quantock Vale	360	2	864	.023
Woolavington	419	2	654	.038
Sowey	257	1	626	.15
Milverton	282	1	544	.17
Islemoor	270	1	524	.17
Castle & Wilton	641	2	448	.074
North Petherton	803	2	381	.098
Glastonbury St Edmunds	383	1	375	.23
Westonzoyland	399	1	373	.24
Galmington	901	2	357	.11
Total MPH Catchment	43751	31	100	-

Table 6 - Electoral Wards with Ten Highest SIR's. Ages 0-24, 1971-1987. MPH catchment used as standard.

WARD	1981 POPULATION (0-24)	NO OF CASES	SIR	P
Holnicote	205	1	833	.11
Quantock Vale	634	2	516	.058
Woolavington	711	2	424	.082
Creech St Michael	750	2	408	.087
Sowey	419	1	396	.22
Williton	822	2	380	.098
Islemoor	441	1	347	.25
Milverton	477	1	339	.26
Glastonbury St Edmunds	490	1	338	.26
Minehead North	589	1	284	.30
Total MPH Catchment	75306	49	100	-

The small numbers involved make detailed analysis of rates difficult, but there appears to be no obvious geographical pattern of incidence. Quantock Vale ward, which contains Hinkley Point, has the highest rate for the 0-14 group, but this is based on just two cases and no conclusion can be drawn from such a small number.

MUSGROVE PARK HOSPITAL CATCHMENT, SOMERSET COUNTY

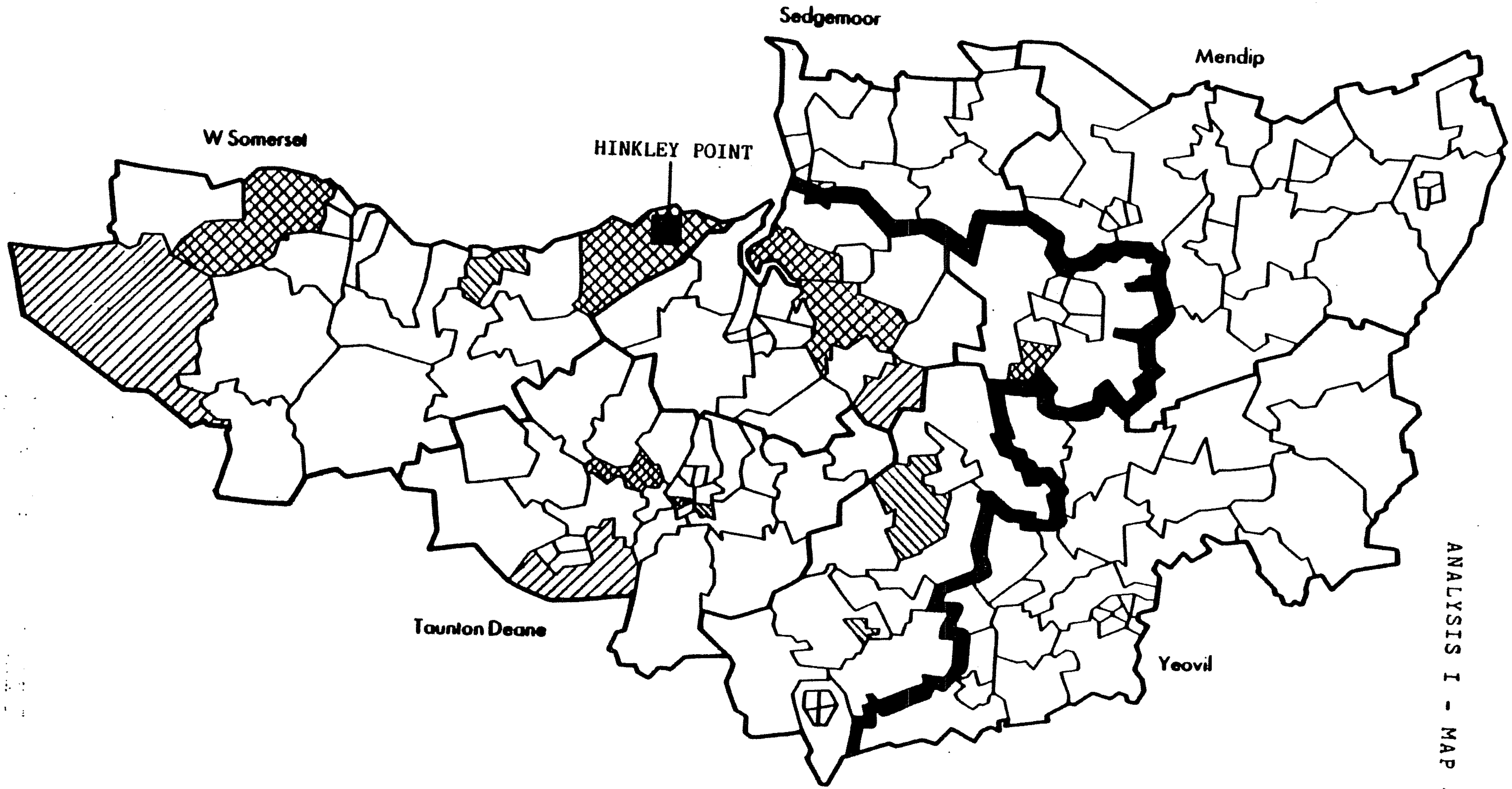


ELECTORAL WARDS WITH HIGH INCIDENCE RATES, 1971-1987

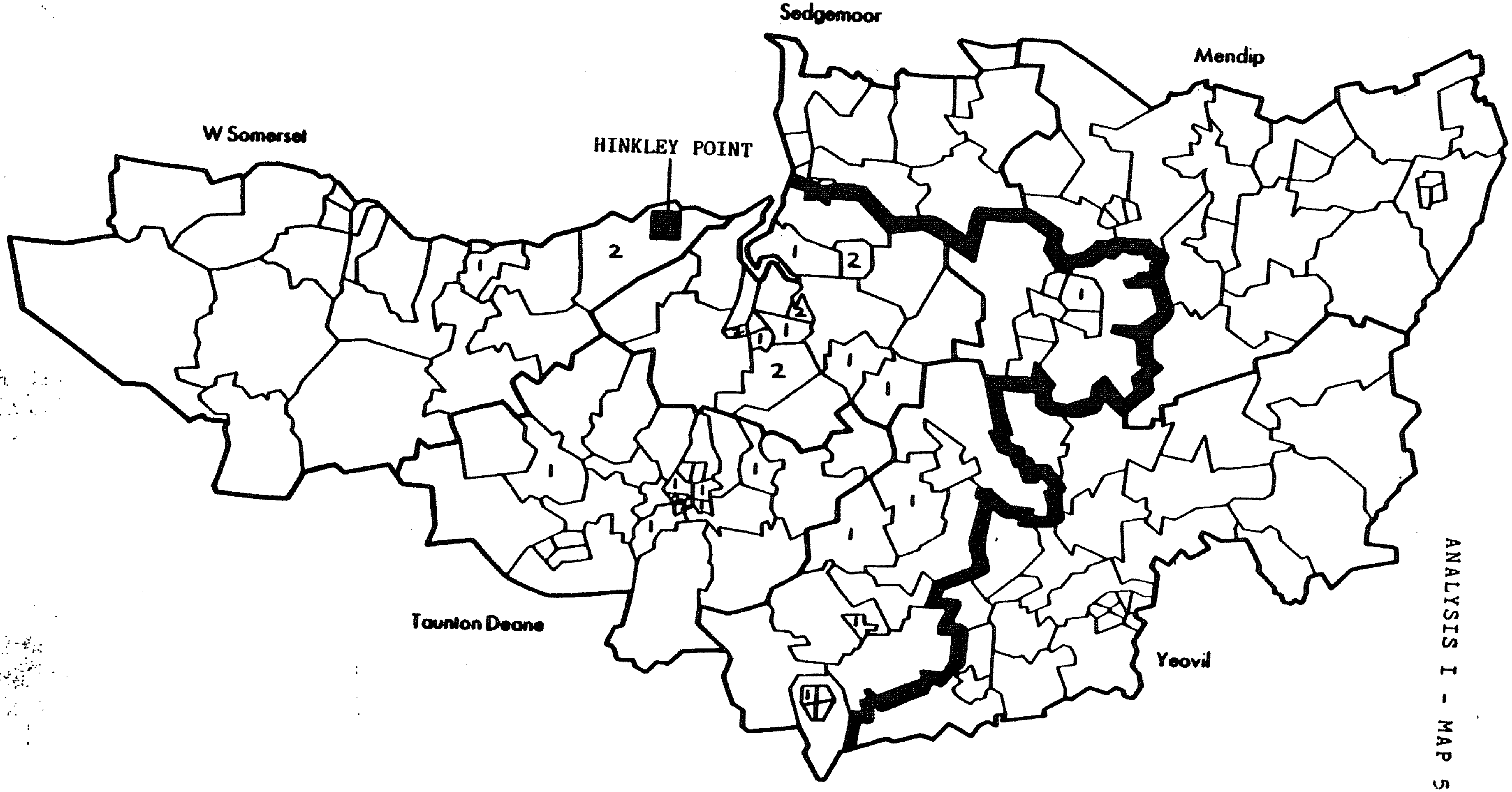
NON-CLL LEUKAEMIAS

Top 10 SIR

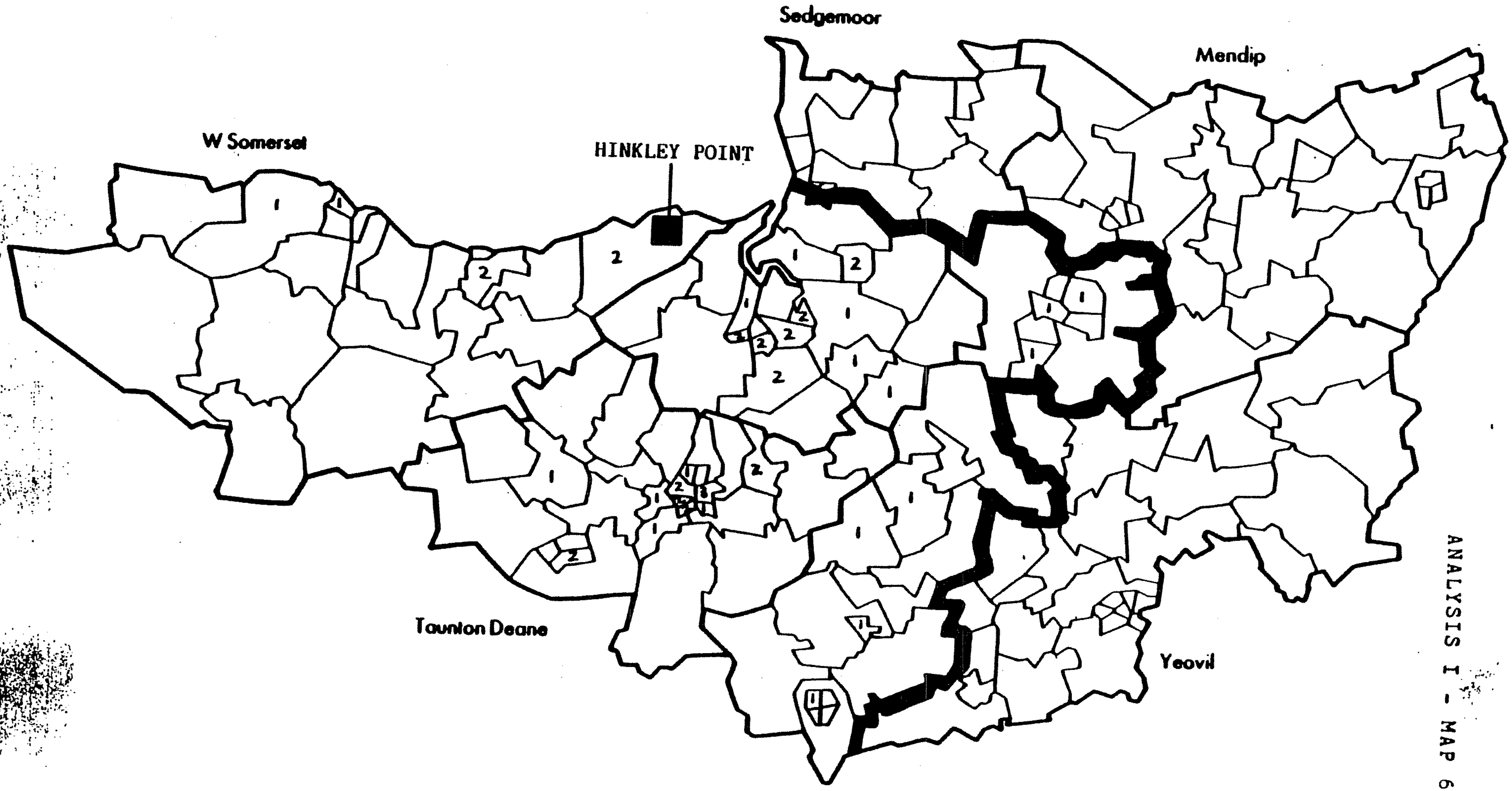
P < 0.05



NUMBERS OF LEUKAEMIAS IN MPH CATCHMENT BY ELECTORAL WARD, 1971-1987. AGES 0-14.



NUMBERS OF LEUKAEMIAS IN MPH CATCHMENT BY ELECTORAL WARD, 1971-1987. AGES 0-24.



ANALYSIS II LEUKAEMIA REGISTRATIONS IN YOUNG PEOPLE IN THE VICINITY OF HINKLEY POINT, 1959-1986

1. Introduction

This analysis aims specifically to test the hypothesis that there is a raised incidence of leukaemia in young persons living in the vicinity of the Hinkley Point nuclear plant. The methodology employed for this purpose follows as closely as possible that used to investigate the incidence of leukaemia around the Dounreay plant. The Dounreay analysis was undertaken initially by the Information Services Division of the Scottish Health Service Common Services Agency[8], followed by a more comprehensive analysis by COMARE[9]. By using the methods used in the Dounreay investigation, a degree of objectivity is assured in that the common problems of how to select geographical boundaries, age limits and methods of case ascertainment are all resolved in advance. Such selection may be crucial to the conclusions drawn. Without specific knowledge of, say, the distance (and direction) from the nuclear plant that radioactive emissions might have an effect, it is difficult to justify the choice of geographic boundary used in the analysis; the choice will be somewhat arbitrary. The problem for the investigator is usually to select a boundary and to convince readers that this decision was made <u>before</u> being aware of the lie of the data. Only when all such decisions are made in advance is the statistical test of an individual hypothesis valid.

This is not to say that variations of the boundaries should not be examined. If a particular hypothesis is repeatedly confirmed choosing different boundaries then the result is somewhat stronger than for a single choice. If, however, the results are particularly sensitive to the choices made then interpretation is more difficult. A problem with multiple hypothesis testing is that it is not surprising to find some tests reaching conventional levels of statistical significance. Thus it is important to have a very limited number of well defined central hypotheses specified in advance.

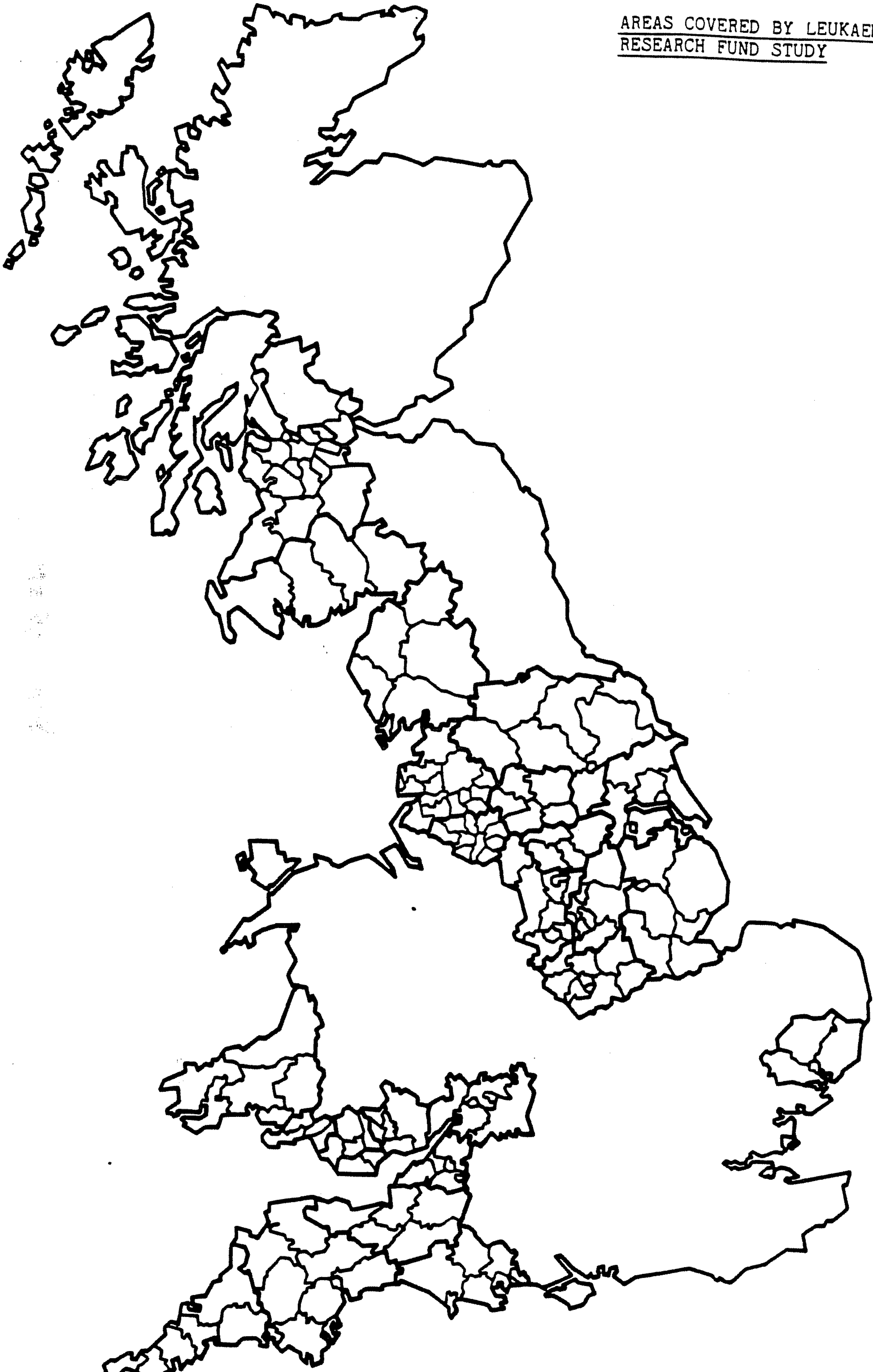
There is no reason why the methods chosen to test the hypothesis in the Dounreay analysis are better or worse than other methods. As mentioned above, the choice of boundaries (geographic, time etc) is often arbitrary. The reason for following the Dounreay methods here is that the precedent has been established. In other words, the choices were made before this study commenced, so there can be no question of selection to fit the data or to suit pre-conceived notions.

2. Case Ascertainment

One of the problems encountered when undertaking the first analysis was finding comparable data to assess the levels of leukaemia in Somerset. Leukaemia cases were established using two sources of information, and comparisons with, say, national rates derived from just Cancer Registry data would not have been valid. Comparable data were found in the Leukaemia Research Fund study, but the data collection for that study only commenced in 1984 and the derived rates may not be valid when applied to earlier years. Mortality data are generally fairly reliable, and Somerset records would almost certainly be fairly compared with national rates. However, with the increasing success in the treatment of leukaemia,

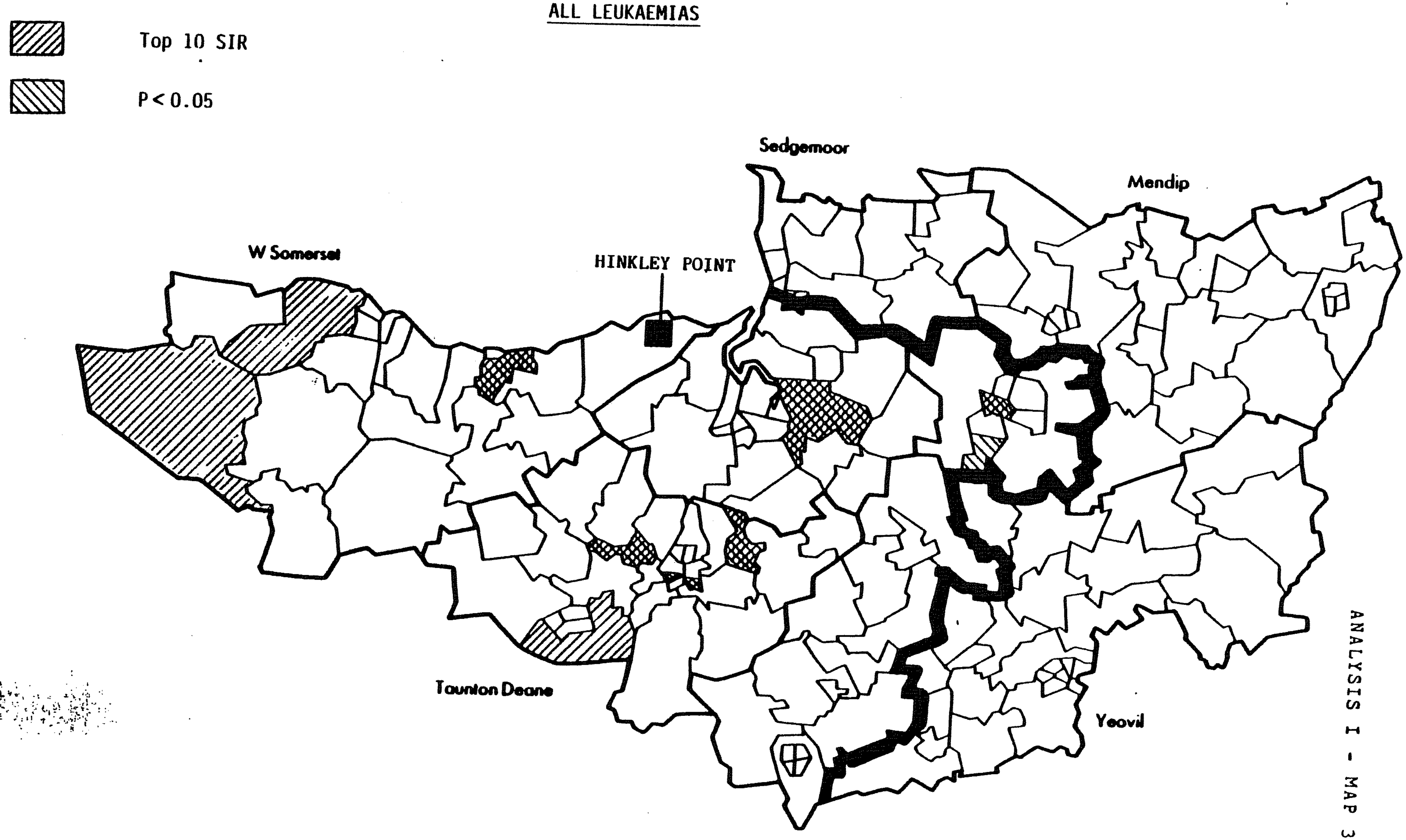
ANALYSIS I - MAP 2

AREAS COVERED BY LEUKAEMIA RESEARCH FUND STUDY



ELECTORAL WARDS WITH HIGH INCIDENCE RATES, 1971-1987

ALL LEUKAEMIAS



mortality records may reflect differential survival patterns as well as incidence.

The COMARE analysis of the Dounreay data used the Scottish National Cancer Registry as its data source. It was, therefore, decided to use the Cancer Registry system in existence in England and Wales for this analysis. Although the number of registered cases for Somerset is less than the number derived when additional information from local records is used (as in the first analysis), comparisons with national registration rates are only valid if the registered cases alone are included.

The South Western Regional Cancer Registry was used to establish the number of cases in the desired age range and disease type (see below), living in Somerset health district at the time of registration, for as many years as practically possible. From computerised and manual records it was possible to obtain sufficient details on all registered cases since 1959.

National rates of cancer registration have been published annually since 1962, initially in the "Registrar General's Statistical Review of England and Wales, Supplements on Cancer", and, since 1971, in the MB1 series of publications by the Office of Population Censuses and Surveys (OPCS) called "Cancer Statistics - Registrations". The most recent such publication deals with 1984 data. Although national registration rates for leukaemias have increased since the early years (possibly due to increasing ascertainment rates rather than incidence), the rates in recent years have not shown any systematic trend, so average rates for the ten year period 1975-1984 were used as an estimate for 1985 and 1986. 1986 was chosen as the end-point of the analysis as it is likely that the 1987 data are incomplete for local cases. Some Registrar General's Cancer Supplements did appear before 1962 but are thought to be unreliable as there was no complete national system in operation. The 1961 volume was, however, published retrospectively based on a complete system. Rates for 1959 and 1960 were estimated using an average of the rates reported for the 10 year period 1961-1970. Again there are no systematic trends evident for individual age/sex groups in the age range considered over those years. Some groups do show a slight increase over the period, probably because the register was less complete in the earlier years. For this reason these estimates for 1959 and 1960 may be slightly high and the Somerset rates may look artificially low by comparison. The effect is not thought to be great but should be borne in mind when viewing the analysis for that period.

3. Validity of the Data

The Cancer Registry relies on voluntary notifications. A scheme for the registration of cancer patients has been in existence since 1945, but a complete national system has only existed since 1962. The national register is simply a collection of independent regional cancer registries, which generally cover health regions.

The validity of cancer registry statistics depends on the completeness of ascertainment and the accuracy of details recorded for each registered person. Since the regional cancer registries have different methods of collecting data, there is likely to be some variation in the degree of validity of the individual registries. In its most recent publication on cancer registrations, OPCS says that "up to 10 per cent of registrations may be missing from the national data". It is likely that this figure masks geographical variations, and for different age and sex groups, cancer types etc, there may be further variations. When analysing time trends, further caution is required as the registration process has undoubtedly improved over the years and an increase in registrations may be reflecting this rather than an increasing incidence.

Various studies have investigated the validity of the cancer registration system for particular regions and/or cancer types, and a recent article reviews and summarises the national perspective[15]. Without special ad hoc studies, a common method to compare regions is to contrast registrations with mortality figures. Variation in registrations/deaths ratios will also reflect differential survival, but should be a reasonable indicator of ascertainment of cancer registries for cancers with low survival rates. Thus OPCS shows for 1984 figures, a registrations/deaths ratio of 1.01 for England and Wales for cancer of the trachea, bronchus and lung in males, with regional figures ranging from 0.97 to 1.10 (with an outlier value of 0.75 for NW Thames which may be partly attributable to changing responsibilities for registration). For breast cancer in females, where survival rates are higher, the national ratio of 1.61 conceals regional figures varying from 1.32 to 1.90. In both cases the South Western ratio was fairly close to the national average (0.97 for male lung and 1.69 for female breast). Calculation of registrations/deaths ratios for leukaemia would not be as helpful because of increasing survival rates for this disease. Comparison of the local records from Musgrove Park Hospital (as described in the first analysis) with Cancer Registry cases, suggests that up to 23% of all known cases may be missing from the cancer registry, although recent years are more complete[14].

Clearly there is considerable variation in the completeness of cancer registries between regions. The implication is that comparison of rates around Hinkley Point with national rates may not be valid if the completeness and accuracy of the South Western registry differs considerably from the national average. One way to avoid this problem would be to use regional rates for comparison rather than national rates. There are two problems with this approach: (i) rates for regions are not routinely available in the same way as national rates; (ii) even within a region there may be different levels of case ascertainment - thus for example Somerset might have different levels of reporting than Cornwall.

Levels of leukaemia for the whole of Somerset health district (see Map 1 at the end of this analysis) are used as well as national rates in this study to compare with those in the immediate vicinity of Hinkley Point. If it is found that rates around Hinkley Point are high compared to national averages, it is important to put this into perspective by reporting Somerset rates. If Somerset rates are similar to national rates then high rates around Hinkley Point are of particular concern. If, however, Somerset rates are generally high with rates around Hinkley Point no higher than these, then Hinkley Point is less likely to be implicated; for example, high rates throughout Somerset may be due to better ascertainment.

4. Age Range Studied

Unlike the first analysis, this section deals only with cancers in young people. It is almost certain that leukaemias arising from exposure to radiation manifest themselves sooner after exposure in children than in adults[4]. It is also likely that this latent period varies less in children than in adults. Thus any increase in leukaemia resulting from general population exposure to radiation over a short period of time is more likely to be detectable in children than in the general population.

It is usual to report cancer data for children based on the age range 0-14. However, COMARE, following the precedent set in the Black report, used the range 0-24; this is partly to obtain sufficient numbers to conduct a proper statistical analysis. This age range is, therefore, used in the Hinkley Point study.

5. Time Periods Studied

As mentioned earlier, local registrations were obtained for 1959-1986. National rates are available for 1961-1984, with those for 1959, 1960, 1985 and 1986 being estimated from the nearest 10 years of available data. Hinkley Point began operations in 1964 with its magnox reactors, the AGR reactors went online in 1977.

If radioactive emissions from Hinkley Point have contributed to extra cases of leukaemia, then higher rates (relative to national figures) would be expected any time after 1964. The questionable completeness of the cancer registry in earlier years and the fact that 1959 and 1960 national rates have to be estimated, make such a comparison of rates before and after 1964 difficult.

The current lack of knowledge regarding the genesis of leukaemia and, in particular, the latent period between exposure to causal agent and onset of disease, makes interpretation of time series data particularly difficult. Nevertheless, the Hinkley Point data are examined over time in a similar way to the analysis of Dounreay data. The number of cases involved is too small to examine annual incidence, so periods of several years are usually constructed for such an analysis. Conventionally, periods of 5 or 10 years centred around census years are taken. The Dounreay analysis followed this convention for 5-year periods as far as possible for the 17 year period considered. In the Hinkley Point study the periods used follow this convention and are, therefore, 1959-1963, 1964-1968, 1969-1973, 1974-1978, 1979-1983 and 1984-1986. The first period coincides conveniently with the years before Hinkley Point was commissioned.

6. Geographical Boundaries Selected

The initial Dounreay analysis examined the level of leukaemia incidence within concentric circles around the nuclear plant. The radii of the circles examined were determined by successively halving an arbitrarily chosen starting distance of 100 km. The two main distances used in the original presentation of the data[8] were 12.5 km and 25 km. Subsequently COMARE, to measure sensitivity to the choice of radii, presented details for other concentric circles. Further halving of the distances were made down to a radius of 3.125 km. In fact, 3.125 km was used as a basic unit to produce all circles from 3.125 km to 25 km in steps of 3.125 km.

Following precedent again, it was decided that the 12.5 km radius should be used for the main test in the Hinkley Point study. A 25 km circle is not used because parts of Wales would need to be included. Furthermore, the area within 12.5 km of Hinkley Point is not so sparsely populated as in the Dounreay case, so that sufficient cases for statistical analysis exist. To analyse sensitivity, it was decided to use circles with radii of 3.125 km, 6.25 km, 9.375 km, 12.5 km and 15.625 km. This follows the COMARE analysis but stops short of wider distances which would complicate matters by the introduction of parts of Avon and Wales. The chosen circles are shown on map 2. It can be seen that the main selected circle, that with radius 12.5 km, intersects the town of Bridgwater, the largest town within the largest circle used (very similar to the problem encountered with the town of Thurso in the Dounreay analysis). The study of other circles is, therefore, particularly useful to test the sensitivity of the chosen 12.5 km. Bridgwater tends to dominate the area in population terms, and its inclusion or exclusion might alter the results considerably.

7. Actual Areas Included

In order to compare local and national incidence rates, it is necessary to have data on the population at risk as well as the number of cases. Thus although it might be possible to determine exactly how many cases occur within 12.5 km of Hinkley Point, it is not known how many people live in such an area, so that rates cannot be calculated. Population figures are generally only available for administrative areas such as parishes, wards, local authorities etc, so that a contortion of the circle is necessary to fit a boundary for which population estimates can be obtained.

Clearly it is desirable to use the smallest administrative areas possible in order to get a reasonable fit to the circle. The smallest areas for which population data are available from the censuses are enumeration districts (ED's), which are constructed specially for each census. In Scotland ED's are coterminous with postcode areas and, since the postcode has been recorded on every case on the Scottish cancer registry since 1968, it was a relatively simple matter for the Dounreay investigators to determine numbers of cases and population at risk for these small areas. In England and Wales, no coterminosity exists between postcode areas and census ED's. Moreover, ED's have not remained constant in the 1961, 1971 and 1981 censuses, especially in areas of new development.

The smallest area that can be used in the Hinkley Point analysis is the rural parish and the urban ward. In rural locations (such as around Hinkley Point) parishes are smaller than wards and are to be preferred. However, in urban areas, wards are the smaller units (in fact civil parishes do not actually exist in a large town, although a "pseudo" parish may be constructed from the whole town). Thus parishes are used here except if a circle transects a town when wards are used. Only the 12.5 km circle passing through Bridgwater requires this refinement.

The five circles were drawn on a parish/ward map and decisions taken as to which parishes and wards should be included in each circle depending on whether the majority of the parish/ward appeared to fall inside. Where the decision was not clear cut on an area basis the actual location of the village(s) in the parish was used to determine whether the majority of the population was likely to lie inside the circle. A map of parishes/wards around Hinkley Point is shown as map 3 while a list of the parishes/wards included in each of the five circles is given in an Appendix at the end of this analysis.

8. Population Estimates

Population figures by sex and five-year age bands for the 1961, 1971 and 1981 censuses are available for parishes and wards in the form of small area statistics (SAS). These are held by the Somerset County Council and were made available for this study.

The national cancer registration rates reported by OPCS are based on the number of cases and the mid-year population estimates. Population estimates are derived from census data but are not equal to aggregations of SAS. For the 1981 census for example, the final mid-year population estimates differ from the SAS in several respects: (1) students and members of the armed forces are allocated to their term-time address or station in the estimates but at their home address in the SAS; (2) wholly absent households on census night are completely excluded from the SAS; (3) an allowance is made for census underenumeration when producing estimates; (4) the estimates are for mid-year and hence allow for three months population change since the census. All these factors contribute to a population estimate for Somerset County in 1981 that is about 3% higher than the SAS figure. For individual parishes and wards (particularly those containing an armed forces station or boarding school), the SAS figure may differ from a population estimate by considerably more (or less) than 3%. However, no official estimates for parishes and wards exist. The 1961 and 1971 censuses differed from the 1981 census in that the population figures were based on "persons present on census night" rather than "persons usually resident".

Although consideration was given to the possibility of producing parish age/sex estimates, it was decided that such a refinement was unrealistic and probably unnecessary. It would be very difficult to produce such estimates as far back as 1959, and it is thought that the discrepancy between SAS and population estimates is unlikely to be a major source of error in the analysis. Thus the SAS from the three censuses were used as the basis for population figures in this study.

There has been no significant change in the boundaries of those parishes and wards around Hinkley Point included in the analysis. Therefore 1961, 1971 and 1981 population figures were produced for each of the five areas by adding together the SAS for the relevant parishes/wards. For consistency the SAS figures were also used for the entire Somerset health district. Figures for years in between the censuses were produced by linear interpolation. 1959 and 1960 figures were produced by backward extrapolation, but for 1982-1986, extrapolation forwards from 1981 was considered too crude; in particular the number of 0-4 year olds decreased between 1971 and 1981 but it is known that this group has increased since 1981. Therefore, the 1981 SAS figures were changed in line with the change in population estimates between 1981 and 1987 to produce comparable figures for the years after 1981. This was achieved directly for the DHA and by using West Somerset and Sedgemoor district council population estimates for the five areas around Hinkley Point (which straddle the two district councils).

9. Types of Cancer Studied

Radiation can play a part in the development of certain cancers, including leukaemia, but there are probably many other causes. The latent period between exposure to radiation and onset of disease appears to be shorter for leukaemia than for other implicated cancers and, as mentioned earlier, this is particularly so for children. Thus leukaemia in children lends itself particularly well to studying possible effects of population exposure to radiation over a relatively short period of time.

There is a disorder closely related to certain types of leukaemia known as non-Hodgkin's lymphoma (NHL) which may also be caused to some extent by radiation. More importantly, increasing refinements in diagnosis have produced criteria for allocation to leukaemia or NHL that are almost certainly different from those used in earlier years. Thus some cases in the 1960's and 1970's labelled NHL might now be called leukaemia, and vice versa. It is therefore sensible to analyse NHL cases as well as leukaemia, and this is done by COMARE. In the original Dounreay analysis the two disease groups studied were "all leukaemias" and "other lymphatic and haematopoietic neoplasms". This latter group contains NHL's but also contains Hodgkin's disease and multiple myeloma. Hodgkin's disease is a separate entity that is not confused with leukaemia, and has not been definitely linked to radiation. Multiple myeloma has been associated with radiation but is not generally found in children (no cases were found in the 0-24 Somerset registrations for 1959-1987). It was therefore decided to analyse NHL's as well as leukaemias, as in the COMARE analysis. The coding system for diseases (International Classification of Diseases (ICD)) has changed twice during the period studied, and the actual ICD codes used are detailed below:-

ICD REVISION	PERIOD COVERED	ICD CODES	
		LEUKAEMIA	NHL
VII	1957-1967	204	200, 202, 205
VIII	1968-1978	204-207	200, 202
IX	1979-	204-208	200, 202

10. Statistical Methods

The usual way of assessing the level of a disease in a particular area is by comparing the observed number of cases (O) with the number expected (E). In the present analysis O is the actual number of cases registered, while E is the expected number based on the annual national rates for each five-year age/sex group and the population size of those groups in each year in the area of interest. The ratio of O to E (usually multiplied by 100 to express as a percentage) is known as the Standardised Registration Ratio (SRR). This is an index of the level of leukaemia found in the area of interest, having standardised for age, sex and year of registration, based on a national average of 100.

The SRR for any one area will rarely equal 100 exactly, even if the disease is randomly spread throughout the population. Natural variation produces some higher values for given areas and time periods and lower values elsewhere. Statistical methods are therefore necessary to assess whether the observed number of cases is consistent with the idea of random disease distribution or whether the data suggest that such random spread is unlikely. In other words, if the observed number of cases in an area is greater than the number expected, how likely is it that the excess number can have occurred simply "by chance"?

Under certain assumptions, the number of cases of a rare disease such as leukaemia in a particular area and time period behaves according to a mathematical law known as the Poisson distribution. The necessary assumptions include equal risk for each individual and cases occurring at random in time and space. The Poisson distribution can be used to estimate how often a particular number of observed cases should occur, given the number expected. A value, P, is calculated which measures the probability of getting as extreme a value of O as the one obtained, for the given E. P is known as the cumulative one-sided Poisson probability and is used as a standard in analyses of this type, including that undertaken by COMARE.

P measures the proportion of times that the observed number of cases (or more) should occur by chance, with very low values of P indicating that the number of cases observed is unlikely to be a chance occurrence. However, no matter how small P is, it can never reach zero, and the possibility of a chance occurrence can never be entirely dismissed. P simply provides a measure of how likely (or unlikely) the occurrence is. Arbitrarily chosen threshold values are often used to determine whether a particular result is "statistically significant". Conventionally the value 0.05 is used but the actual P value is more important than whether it happens to be lower than some arbitrary value.

If the number of leukaemias and/or NHL's within a given distance from the power station is large, the P value associated with that number of cases gives an indication of how unlikely it is to have occurred by chance. It has been argued in this context that leukaemias can occur in "clusters" anyway and that the number of cases found near the plant is not so uncommon as the P value suggests. In other words, the assumptions underlying the use of the Poisson distribution are called into question (in particular that the disease is randomly spread throughout the population with everyone being at equal risk) . There is still debate about whether leukaemias actually are randomly spread throughout the population. A recent study of leukaemia mortality throughout England and Wales found variation in excess of that expected from Poisson theory (Sarah Darby, personal communication). In the COMARE analysis, the Poisson assumption was verified by constructing areas throughout Scotland of roughly the same population size as the 12.5 km and 25 km circles around Dounreay and examining whether the numbers of cases in each of these areas conformed to a Poisson distribution. The actual numbers observed were close to the numbers predicted and it was concluded that there was no significant deviation from the Poisson model. Similarly, in the Black report, the distribution of leukaemias in rural districts throughout England and Wales similar in size to Millom rural district (which was under scrutiny because of its closeness to Sellafield) was examined and found to conform to a Poisson distribution. These checks give validity to the use of the Poisson distribution in assessing how likely the number of cases around the nuclear plants are to have occurred by chance.

In the present analysis it has not been possible specifically to assess the appropriateness of using the Poisson distribution to analyse the numbers of cases around Hinkley Point. A special algorithm was used in the COMARE analysis to generate the artificial areas for analysis. This algorithm would not be applicable in the Somerset case because the lowest geographical level analysed here is parish (and ward) and it would be difficult to construct areas of similar population size to those studied around Hinkley Point. However, the confirmation of the use of the Poisson distribution as a reasonable model in the Black and COMARE reports gives some support to its use here. The two reports demonstrate that the model is satisfactory for rural districts in England and Wales and for areas in Scotland with the same population sizes as those studied around Dounreay. There is no reason to suppose that the number of leukaemias in the area around Hinkley Point should behave any differently, i.e. other than according to the Poisson law.

11. Registered Cases of Leukaemia and NHL

A total of 22 leukaemias and NHL's in persons aged under 25 have been registered during 1959-1986, with an address at time of registration falling in the group of parishes within the 12.5 km circle around Hinkley Point nuclear power station. Two of the leukaemias and one NHL were registered before Hinkley Point commenced operations in 1964. A full list of the cases is shown in table 7.

Table 7 - Leukaemia and NHL Registrations, 1959-1986, Age 0-24
Address within 12.5 km of Hinkley Point

Case No.	Year of Registration	Year of Birth	Age at Registration	Sex	Disease Type	Approximate distance H.Pt. (km)
1	1960	1938	22	F	NHL	12
2	1962	1944	18	M	Leuk	12
3	1962	1944	18	F	Leuk	12
4	1964	1945	19	M	NHL	10
5	1964	1940	24	M	Leuk	12
6	1964	1948	16	F	Leuk	10
7	1965	1962	3	F	Leuk	11
8	1966	1961	5	M	Leuk	12
9	1969	1952	16	F	Leuk	7
10	1969	1945	24	F	Leuk	7
11	1970	1969	1	F	Leuk	12
12	1970	1962	8	F	Leuk	7
13	1971	1968	3	M	Leuk	12
14	1971	1950	21	F	NHL	12
15	1972	1959	13	F	Leuk	3
16	1973	1968	5	F	NHL	8
17	1973	1958	15	M	Leuk	12
18	1977	1971	6	M	NHL	10
19	1979	1961	18	M	Leuk	11
20	1980	1973	7	M	Leuk	12
21	1983	1964	19	M	NHL	10
22	1984	1962	22	F	NHL	11

Cases of leukaemia and NHL have occurred throughout the entire study period, but relatively few cases have occurred since 1973. There appears to be a heavy concentration of cases between 1969 and 1973, with a similar concentration during 1964-1966. The ages of the cases registered also cover the entire range under study.

12. Leukaemia and NHL Incidence - Main Result

As described earlier, the number of cases around Hinkley Point can be compared with the number expected from national rates. For comparison purposes, the incidence for the remainder of the Somerset Health District (SHD) is also examined.

The main result relates to the numbers of cases inside the 12.5 km boundary registered since 1964 (i.e. when Hinkley Point was commissioned); table 8 gives the details.

Table 8 - Leukaemia and NHL Registrations 1964-1986. Age 0-24

	1981 Population	Observed (O)	Expected (E)	SRR (100 x O/E)	P
Leukaemias					
< 12.5 km	12670	13	7.91	164	0.060
SHD residue	112700	87	75.56	115	0.11
SHD total	125370	100	83.47	120	0.043
NHL's					
< 12.5 km	12670	6	2.51	239	0.043
SHD residue	112700	31	24.37	127	0.11
SHD total	125370	37	26.88	138	0.037
Leukaemias & NHL's					
< 12.5 km	12670	19	10.42	182	0.011
SHD residue	112700	118	99.93	118	0.042
SHD total	125370	137	110.35	122	0.0079

The results show that for leukaemias and NHL's (as well as both together) there are significantly more cases in Somerset than might be expected from national rates, using the conventional 5% level of significance. In addition, the SRR's are higher for the area around Hinkley Point than for the remainder of SHD. For leukaemias and NHL's separately the SHD residue does not have statistically significantly higher rates than the national average. The 12.5 km circle around Hinkley Point has a leukaemia rate that just fails to reach the conventional level of significance and a NHL rate that is just significant. For both leukaemias and NHL's together, the rate for SHD residue is just significant while for the area around Hinkley Point there were 82% more cases than expected, with an associated P value of 0.011, i.e. such an observation should only occur by chance about 1.1% of the time.

Taken in isolation, the rates around Hinkley Point show a significantly higher incidence in that area than might be expected from national rates. However, the fact that the remainder of SHD also has high rates makes interpretation difficult. As SHD has high rates generally, it can be argued that high rates around Hinkley Point are not surprising; for example it could be due to better reporting throughout Somerset. On the other hand, it might be argued that the "effect" of Hinkley Point is more widespread than 12.5 km and that is why the SHD residue is fairly high compared to national rates (but not as high as rates inside the 12.5 km range). To answer this last point would be particularly difficult; South Western Regional rates could be examined for comparison, but then two more nuclear power stations (at Oldbury and Berkeley) would be included and the issues still more confused.

It can be seen from table 8 that although the SHD residue has significantly high rates for leukaemias and NHL's together, the SRR in absolute terms is not particularly high - 118 compared to a national average of 100. The SRR for the area around Hinkley Point is considerably higher at 182; similar observations can be made for leukaemias and NHL's separately. An obvious question to ask is whether the SRR for the 12.5 km area is statistically significantly higher than the SHD residue. This hypothesis can be tested statistically using knowledge about the behaviour of two independent Poisson variables. Conditional on the <u>total</u> observed number of cases throughout SHD, the observed number in the 12.5 km zone should behave according to a formula known as the binomial distribution. The two parameters of this distribution are the "sample size" (which here is the total number of observed cases), and the probability of a case occurring in the 12.5 km zone (which is the expected number in the zone divided by the total expected number throughout SHD). It can be shown by exact calculation of the binomial probability that, given there are a total of 137 cases in SHD as a whole, the probability of 19 or more being in the 12.5 km area is 0.058. That is, the difference between the SRR for the 12.5 km zone and that for the SHD residue just fails to reach the conventional 5% level of statistical significance.

13. Sensitivity to Choice of Geographical Boundary

Although the 12.5 km boundary was selected before inspecting the data, it is useful to see how sensitive the results are to this choice. Thus other boundaries selected in advance are used in this section to examine the sensitivity. Table 9 presents results for circles around Hinkley Point with radii calculated as multiples of 3.125 km as described earlier. Note that the SHD residue here is that part of SHD that is outside the 15.625 km circle and differs from that used in table 8.

For leukaemias alone, the SRR for every circle around Hinkley Point is higher than that for the SHD residue. For NHL's there are no cases inside the smallest two circles but for larger circles the SRR is again higher than SHD residue. For leukaemias and NHL's combined, all areas around Hinkley Point produce higher SRR's than the SHD residue figure. However, because of the small numbers involved, conventional levels of statistical significance are not always reached. The 5% level is in fact never reached for leukaemias alone, the 12.5 km zone producing the smallest P value of 0.06. For NHL's the largest two circles around Hinkley Point have significantly high rates. For leukaemias and NHL's combined, the one case within 6.25 km fails to produce significantly high rates for the two smallest zones. The next largest zone just fails to reach statistical significance. As reported in the previous section, the 12.5 km zone reaches significance with a P value of 0.011. The extra result here is that this level of significance is maintained for the larger circle of radius 15.625 km. Moreover the rate for the SHD residue outside this circle now fails to reach statistical significance. Using the statistical test described above, it can be shown that the largest zone around Hinkley Point does not have statistically significantly higher rates than the SHD residue (P = 0.090), so the problem of interpretation remains.

Table 9 - Leukaemia and NHL incidence, 1964-1986. Age 0-24

	1981 Population	Observed (O)	Expected (E)	SRR (100 x O/E)	P
Leukaemias					
< 3.125 km	420	1	0.310	323	.27
< 6.25 km	840	1	0.570	175	.43
< 9.375 km	2440	4	1.71	234	.095
< 12.5 km	12670	13	7.91	164	.060
< 15.625 km	25130	22	16.04	137	.091
SHD residue	100240	78	67.43	116	.11
SHD total	125370	100	83.47	120	.043
NHL's					
< 3.125 km	420	0	0.0975	0	1
< 6.25 km	840	0	0.177	0	1
< 9.375 km	2440	1	0.549	182	.142
< 12.5 km	12670	6	2.51	239	.043
< 15.625 km	25130	11	5.10	215	.016
SHD residue	100240	26	21.78	119	.209
SHD total	125370	37	26.88	138	.037
Leukaemias & NHL's					
< 3.125 km	420	1	0.407	246	.33
< 6.25 km	840	1	0.747	134	.53
< 9.375 km	2440	5	2.26	221	.079
< 12.5 km	12670	19	10.42	182	.011
< 15.625 km	25130	33	21.15	156	.010
SHD residue	100240	104	89.20	117	.068
SHD total	125370	137	110.35	122	.0079

14. Time Trends

The time periods selected for study are conventional 5-year periods centred around census years, except for the most recent 3-year period. Results for the 12.5 km zone and the rest of the SHD are shown for these periods in table 10. Figures for the 5-year period before the commissioning of Hinkley Point are included, as well as a breakdown of the entire period studied in previous sections.

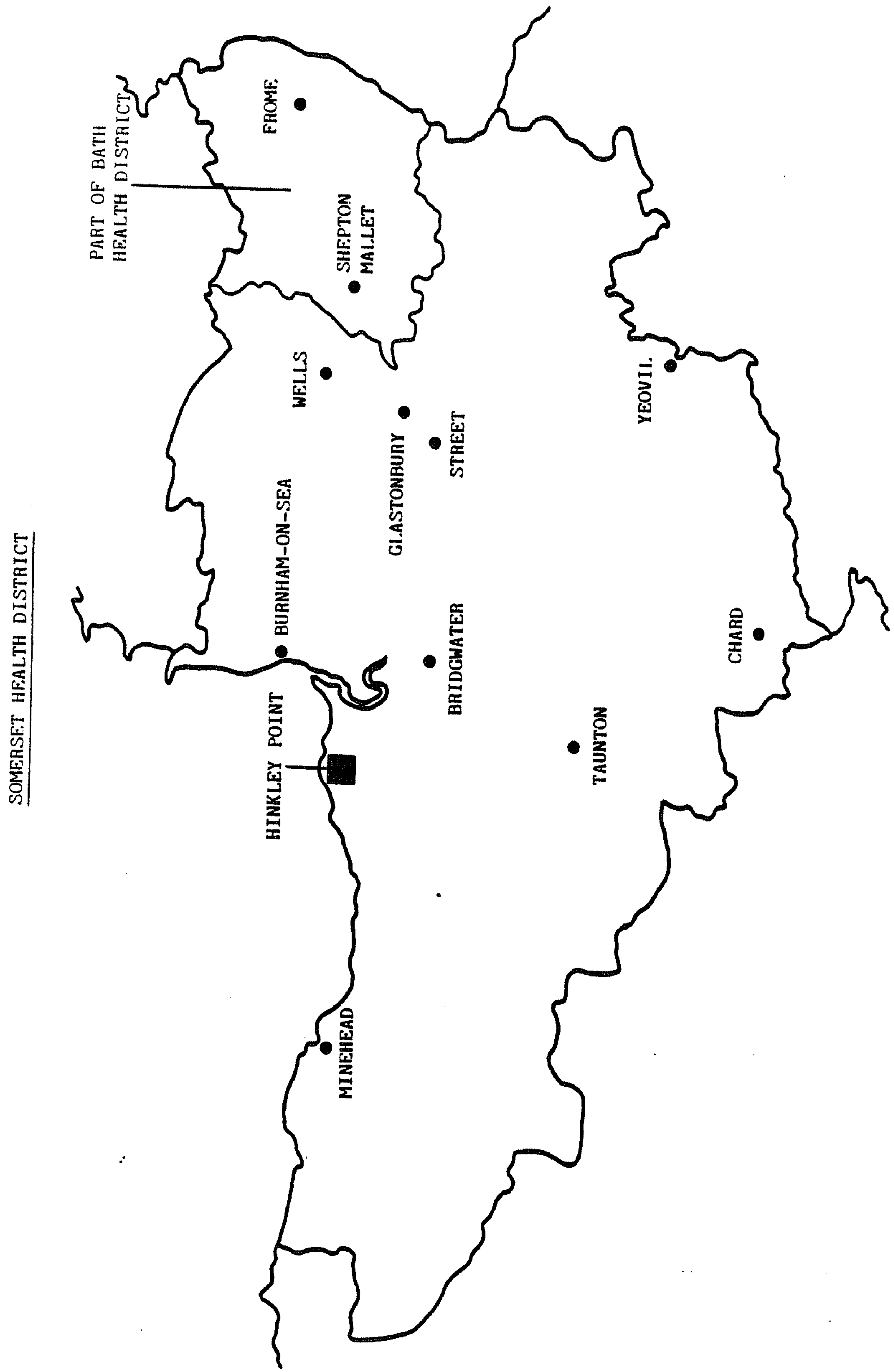
It is evident from the table that the excess cases within the 12.5 km zone are concentrated in the ten year period 1964-1973. The two 5-year periods within this have statistically significantly high rates, with the 1969-1973 period in particular having a very high SRR; in this 5-year period there were about four times as many cases of leukaemia and NHL than would be expected from national rates during the same years (O=9, E=2.27) and the P value suggests that such an occurrence should happen by chance on about 1 in 1700 occasions. The rates for the rest of SHD during 1964-1973 are very close to the national rates, so there would seem to be no question of the high rates around Hinkley Point simply reflecting high rates in Somerset generally.

None of the subsequent time periods produce significantly high rates for the 12.5 km zone. SHD residue rates are higher than national rates but not significantly so. The only time period for which the SHD residue rate is significantly high is 1959-1963 (before Hinkley Point commenced operation); in this earliest period the 12.5 km zone had a high rate but not significantly so.

The results presented in this section show that the number of leukaemias and NHL's near Hinkley Point within ten years of the power station commencing operation, is significantly higher than the number of cases expected from national rates, at a time when the rest of Somerset did not have high rates. Rates for subsequent years are unremarkable. The 5-year period 1969-1973 shows particularly high rates. It is noteworthy that all five cases reported in table 9 as falling within 9.375 km of Hinkley Point occurred in this time period. However, the details of individual cases given in table 7 shows that the five cases are spread throughout the 5-year period and are of differing ages.

Table 10 - Leukaemia and NHL Registrations, Age 0-24

	Leukaemias			NHL's			Leukaemias + NHL's		
	O	SRR	P	O	SRR	P	O	SRR	P
<u>1959-63</u>									
< 12.5 km	2	159	.36	1	209	.38	3	173	.25
SHD residue	19	147	.07	10	199	.03	29	162	.0099
<u>1964-68</u>									
< 12.5 km	4	296	.048	1	175	.44	5	260	.046
SHD residue	17	126	.20	3	52	.93	20	104	.47
<u>1969-73</u>									
< 12.5 km	7	401	.0022	2	381	.10	9	396	.00059
SHD residue	16	94	.63	5	96	.60	21	94	.63
<u>1974-78</u>									
< 12.5 km	0	0	1	1	193	.40	1	41	.92
SHD residue	24	119	.11	10	204	.029	34	147	.20
<u>1979-83</u>									
< 12.5 km	2	114	.53	1	178	.43	3	129	.41
SHD residue	17	107	.43	10	193	.038	27	128	.12
<u>1984-86</u>									
< 12.5 km	0	0	1	1	300	.28	1	69	.76
SHD residue	13	120	.30	3	91	.64	16	113	.34


SOMERSET HEALTH DISTRICT
PART OF BATH HEALTH DISTRICT
FROME
SHEPTON MALLET
WELLS
GLASTONBURY
STREET
YEOVIL
BURNHAM-ON-SEA
BRIDGWATER
CHARD
HINKLEY POINT
TAUNTON
MINEHEAD

DISCUSSION

1. Why Two Analyses?

At first sight there may be some confusion for the reader, confronted with the results of two separate analyses. Indeed, a single study would have been preferable if it had been possible to use the superior aspects from each of the two methodologies, eliminating where possible the relative weaknesses.

Such a course is not possible. Analysis II specifically tests a hypothesis relating to Hinkley Point, and the approach is to be preferred if that is the only interest. However, the cancer registry, known to have inadequacies, was the only source of data used in this analysis. Analysis I, on the other hand, uses data collected in a more efficient manner, and the ideal would be to conduct the second analysis using such data. There are three main reasons why this can not be done:

(i) The 12.5 km circle around Hinkley Point includes an area that is not within the nominal catchment of MPH, namely Burnham-on-Sea and surrounds. Although some cases from this area are referred to MPH, some will be referred to Bristol and it is not possible to obtain accurate information on these cases.

(ii) Local records on lymphoma cases have not been maintained to the same extent as those on leukaemia cases.

(iii) The system of maintaining a relatively complete record of all leukaemias only commenced in 1971. It is clearly desirable to analyse incidence since Hinkley Point commenced operation.

2. Summary of Results

A summary of the findings in the two analyses is as follows:-

Analysis I

Rates of leukaemia incidence for all ages in the whole of the MPH catchment are high compared to the study areas used by the LRF. The excess is 24% over the period 1971-1987, but if the comparison is limited to just that period for which LRF rates are available (1984-1986) the excess is 61%.

Electoral wards with high rates over the 1971-1987 period appear to be scattered throughout the catchment, with no evidence of "clustering" around Hinkley Point. The ward containing the power station does have a high rate for radiation related leukaemias, but not exceptionally so in the context of other wards in the MPH catchment.

Childhood leukaemias also are scattered throughout the catchment. For the 0-14 age group the ward containing Hinkley Point has the highest rate of the 91 wards, but this is based on just two cases.

In general, the conclusions relating to this analysis agree with those presented in the previous report. There is no evidence of clustering of leukaemia cases in the MPH catchment; in particular the rates around

APPENDIX

PARISHES AND WARDS WITHIN 15.625 KM OF HINKLEY POINT

Within 3.125 km

Stogursey

Within 6.25 km

As above, plus:

Fiddington
Otterhampton
Stockland Bristol
Stringston

Within 9.375 km

As above, plus:

Cannington
East Quantoxhead
Holford
Kilve
Nether Stowey
Over Stowey
Pawlett

Within 12.5 km

As above, plus:

Berrow
Bicknoller
Burnham-on-Sea and Highbridge
Chilton Trinity
Crowcombe
Enmore
Puriton
Quantock (Bridgwater ward)
Spaxton
Victoria (Bridgwater ward)
Wembdon
West Huntspill
West Quantoxhead

Within 15.625 km

As above, plus:

Bawdrip
Brean
Brent Knoll
Bridgwater Without
Broomfield
Burnham Without
Central (Bridgwater ward)
Chedzoy
Cothelstone
East Huntspill
Eastover (Bridgwater ward)
Goathurst
Hamp (Bridgwater ward)
Lympsham
Sampford Brett
Stogumber
Sydenham (Bridgwater ward)
Watchet
West Bagborough
Williton
Woolavington

Hinkley Point do not of themselves give cause for concern. It should be noted, however, that this analysis did not set out specifically to test a hypothesis relating to Hinkley Point.

Analysis II

For the period since Hinkley Point commenced operations (1964-1986), the number of leukaemias and lymphomas in persons aged under 25 within 12.5 km of the power station, registered with the South Western Cancer Registry, is significantly higher than the number expected from national rates. However, this result must be examined in the context of rates for the rest of Somerset; the high rates around Hinkley Point might simply reflect high registration rates throughout Somerset.

The rates for the rest of Somerset are indeed high compared to national rates. Nevertheless the rates around Hinkley Point are higher still, although comparisons with the Somerset residue fail to reach conventional levels of statistical significance. The possibility that the high rate around Hinkley Point may simply reflect natural variation cannot, therefore, be excluded.

Examination of pre-determined time periods reveals that the excess numbers of registered cases near Hinkley Point are concentrated in a 10-year period following the commissioning of the power station. In this period the rest of Somerset experienced rates close to the national rates. The period 1969-1973 shows particularly high rates (9 cases when 2.27 were expected), with an estimated 1 in 1700 probability of this being a chance occurrence. There was, however, no reason to suspect in advance this particular period, and the problem of multiple testing must be borne in mind; ie, if many statistical tests are conducted, the probability of achieving statistical significance on at least one is considerably higher than the nominal testing level. Nevertheless, the P value obtained here suggests that the high rate found is unlikely to be a chance occurrence.

3. Limitations of the Analyses

As has been described throughout the report, there are several potential sources of error and limitations in analyses of the type undertaken. Each of the two separate analyses has its own weaknesses (and strengths) and some limitations are common to both analyses. Such limitations must be recognised when attempting to interpret the results; some of the problems associated with interpretation are now described.

Specific Limitations of Analysis I

(i) The main limitation of the first analysis is that it did not set out to test a specific hypothesis. For reasons described above, it is not possible directly to test a hypothesis about leukaemias in the vicinity of Hinkley Point using the data collection methods of the first analysis.

(ii) Only data for the catchment of Musgrove Park Hospital are examined. In particular, Burnham-on-Sea, a town along the coast in the line of prevailing winds from Hinkley Point, is excluded.

↗ N.B. Also Weston-S. Mare

(iii) Only data since 1971 are available. Hinkley Point commenced operations in 1964 and it is desirable to examine the entire period since then. (This is particularly relevant in the light of the finding of the second analysis, that high rates in the vicinity of Hinkley Point seem to be concentrated in the period 1964-1973.)

(iv) Data on lymphomas are not available. The criteria for assigning particular types of leukaemia and lymphoma have changed over time and some cases in earlier years might possibly be re-assigned according to current criteria.

(v) The LRF rates are for 1984-1986 only. Such rates are almost certainly not applicable to earlier years, since there appears to be an increase in incidence over time; this problem of comparability remains even if the explanation for the increase is better case ascertainment.

(vi) No comparison with external rates for childhood leukaemias is possible.

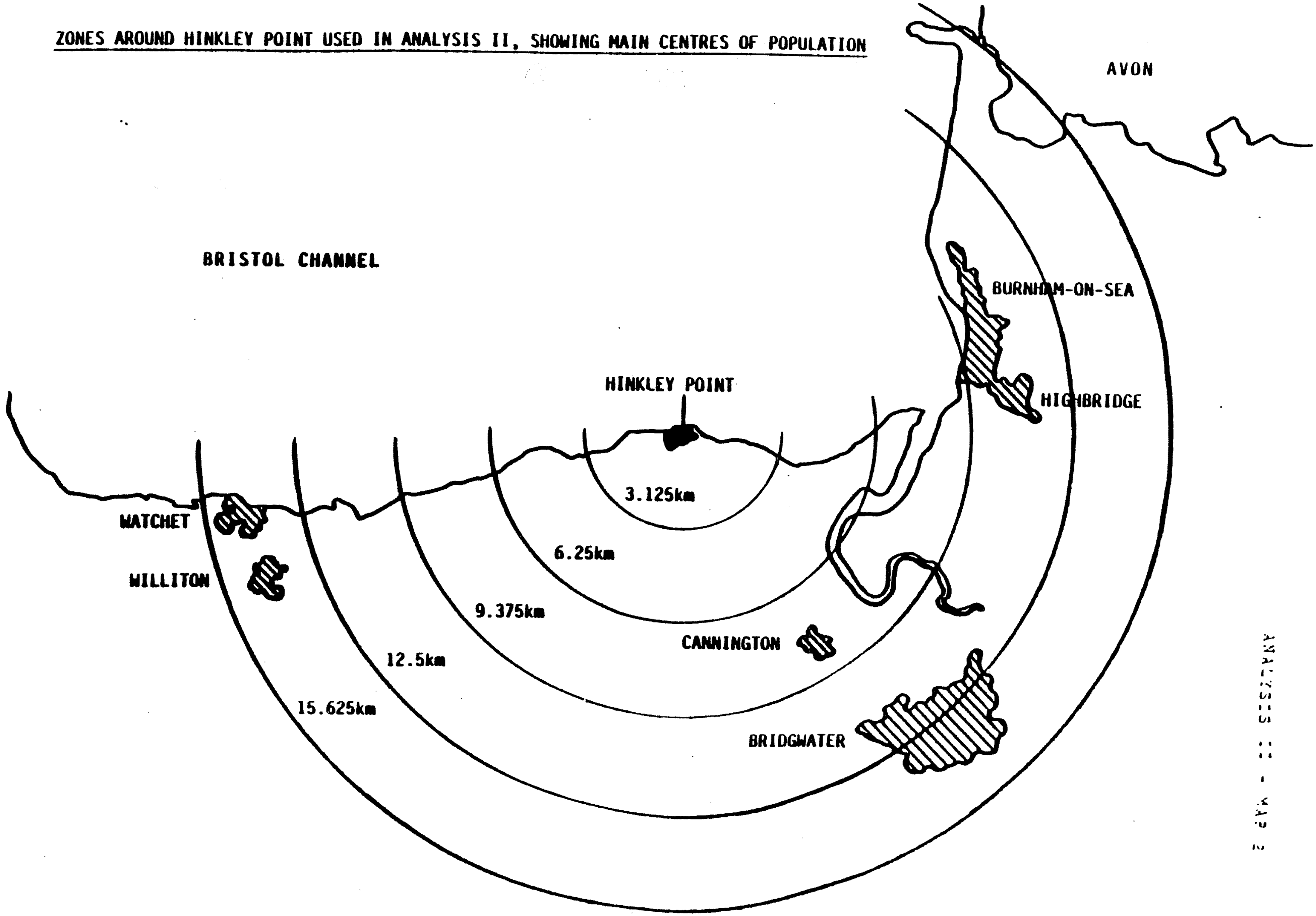
Specific Limitations of Analysis II

(i) The important limitation of the second analysis (and the whole reason why the first analysis was conducted) is that the cancer registry does not provide a complete assessment of leukaemia incidence. Thus local cases are missed, but national rates also are based on incomplete ascertainment. It is not possible to say for sure whether Somerset is better at reporting cases than elsewhere. The first analysis identified 48 cases of leukaemia in the under-25 age group in the MPH catchment since 1971; four of these are missing from the cancer registry (none of these is inside the 12.5 km zone but one is inside the 15.625 km zone).

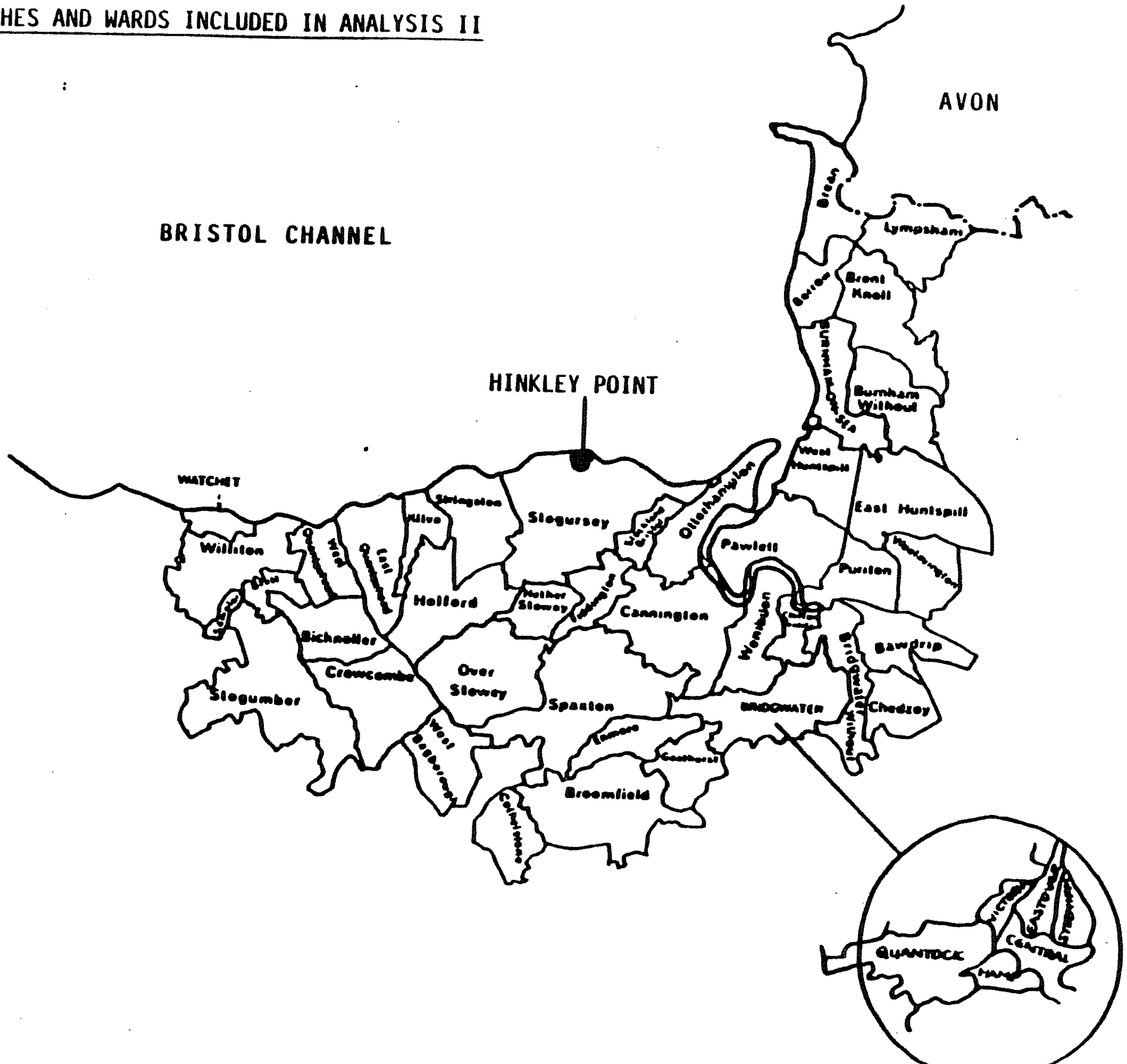
(ii) The selection of geographical zones is somewhat arbitrary. If radioactive emissions from Hinkley Point contribute to extra cases of leukaemia, it is impossible to know in advance which areas are likely to be affected. It might be argued, for example, that direction of prevailing winds should be taken into account or perhaps tidal movements. In the absence of appropriate knowledge, there is no reason to suppose that the 12.5km circular zone is any worse than any other choice. The precedent has been set in the Dounreay analysis and the methods replicated here. The sensitivity analysis conducted suggests that the results found are not peculiar to the choice of the 12.5km boundary.

(iii) The selection of time periods is arbitrary. The excess of leukaemias discovered in the analysis was found to be concentrated in a 10-year period, and in particular the 5-year period 1979-1983. Although the time periods were selected in advance, there was no reason to suspect any particular period as more likely to have increased rates. Natural variation is bound to produce some time periods with higher rates than others, and statistical tests on several time periods are quite likely to produce at least one significant result. However, the size of the P value in this particular case suggests that the possibility of the result being a chance occurrence is unlikely.

ZONES AROUND HINKLEY POINT USED IN ANALYSIS II, SHOWING MAIN CENTRES OF POPULATION



PARISHES AND WARDS INCLUDED IN ANALYSIS II



Limitations common to both analyses

(i) The numbers of cases involved is very small. Interpretation of excess numbers of cases is very difficult when the absolute number of extra cases is small. The best example of this problem is in analysis I where, for the under 25 year olds, the electoral ward with the highest SIR has a rate more than eight times that expected, but this is based on just one case. Any errors involved, such as a misrecorded diagnosis, will have great impact because of the small numbers.

(ii) Neither analysis takes account of the mobility of the population. Address at time of diagnosis has been used in the geographical definition, and no account has been taken of movement in and out of the area. Some cases may, for example, have recently moved into the area, making any link with Hinkley Point unlikely. There may be cases diagnosed elsewhere that have recently moved out of Somerset; the disease may have been initiated while in Somerset with clinical symptoms appearing later.

(iii) The analyses do not provide any evidence of cause and effect. A finding of excess cases of leukaemias near a nuclear power plant does not necessarily imply that the excess is caused by the station. Such a finding for several power stations makes it more likely that the stations are implicated in some way, but not necessarily because of radioactive discharges.

4. Possible Explanations of Results

The second analysis has revealed an unusual pattern of leukaemia incidence in young people living near Hinkley Point. The first analysis does not particularly support this finding, but neither does it contradict it; it was not designed specifically to test the hypothesis. Moreover, the fact that the high rates are confined to the period 1964-1973 makes it unsurprising that the first analysis does not produce corroborative evidence, as it only covers the period since 1971.

The reason for the unusual pattern of incidence is not known. The final conclusion of the COMARE report stated that the Dounreay data, taken in conjunction with the Sellafield data, seemed to indicate that some feature of those two plants "...leads to an increased risk of leukaemia...". Bearing in mind that the second analysis of this present report follows as closely as possible the analysis of the Dounreay data, it would seem fair to view the results here in conjunction with the results in the COMARE report. In that case it becomes increasingly persuasive that some feature of the nuclear plants is indeed leading to raised levels of leukaemia. However, it should be noted that the temporal pattern of the results for the Hinkley Point area contrasts sharply with that for the Dounreay area, where the excess cases of leukaemia were concentrated in the most recent period examined (1979-1984). Furthermore, the patterns of leukaemia incidence do not in themselves imply that radioactive emissions from the stations must be causing the high rates. Nuclear power stations have other things in common (apart from their radioactive discharges), and may be associated with high leukaemia incidence for a totally different reason. Any plausible explanation in the local context should encompass the high rates in the 1969-1973 period and the lower rates found in subsequent years.

Possible explanations for the results in this report need to be examined. Several possibilities are described in this section, some relevant to the local context only, others applicable to nuclear power stations generally.

(i) The results relating to Hinkley Point could be a chance finding. Although rates around Hinkley Point are statistically significantly higher than national rates for the entire period since the station commenced operations, they are not significantly higher than rates for the rest of Somerset, using the conventional 5% level of significance. Rates for Somerset as a whole are high (confirmed by both analyses) and indeed they were high in the five-year period before Hinkley Point was commissioned. It is not clear why rates are high throughout Somerset. One possibility is that the social class distribution in Somerset lends itself to high rates. A recent study shows that areas in England and Wales with a high social class structure tend to have relatively high rates of mortality from leuakeamia, although the reason is unknown (Sarah Darby, personal communication).

In the context of Somerset rates, the incidence around Hinkley Point is high but not exceptional for the entire period 1964-1986. However, the rates during 1964-1973 are high around Hinkley Point at a time when rates for the rest of Somerset were close to the national figures. In particular, the rates for the period 1969-1973 are unlikely to have arisen by chance.

(ii) Young people in the vicinity of Hinkley Point could have been in contact with some unknown cancer-producing chemical or toxin. There is, for example, a number of factories in the area.

(iii) The excess cases could have been caused by a virus infection. Some types of leukaemia have a virus implicated in their aetiology. The building of a large complex like Hinkley Point requires a large work force, and new families brought together with different levels of immunity are particularly susceptible to virus infections. Infection by a leukaemia-producing virus in a few susceptible children could initiate the disease.

A recent study[17] has given some credence to this theory. The study examined the incidence of leukaemia in young people living in a remote rural area, subject to a large influx of people over a relatively short period of time. The situation was thought to resemble that when large numbers of people come together to build a nuclear installation, but in this case no such installation was being built. A significant increase of leukaemia was found in the study area, and it was argued that this was suggestive of the leukaemias arising as a rare response to a common virus infection. This theory, if true, could explain at least some of the excess leukaemia cases found around some nuclear installations.

(iv) Radioactive emissions from Hinkley Point could have caused the extra cases of leukaemia. The particular temporal pattern of the cases suggests that, if this is the case, then a large unreported release must have occurred in the 1960's. Although a detailed assessment has not been possible it is likely that such a release would need to have been thousands of times the reported annual releases.

5. Conclusion

There is evidence of a raised incidence of leukaemia and non-Hodgkin's lymphoma in young people living in the vicinity of the Hinkley Point nuclear power station. In particular, there are high rates in the ten-year period following Hinkley Point commencing operations. No single cause for this finding is identified. Epidemiological investigations into possible common factors of the cases occurring during the ten-year period need to be undertaken to seek to discover causes. The possibility of a major unreported release of radioactive material in the 1960's from the nuclear power station needs to be explored.

ACKNOWLEDGEMENTS

Many people have helped in making this report possible. It should be noted, however, that the final interpretation and views expressed in the report are those of the authors; any errors are theirs alone.

Specific thanks are due to:

- Dr Ray Cartwright (Director) and other staff at the Leukaemia Reseach Fund Centre for Clinical Epidemiology, for providing "expected" numbers of cases based on their data collection survey, and for general advice.
- Drs Malcolm Phillips and Steve Johnson, consultant haematologists at Musgrove Park Hospital, for use of their local register, help in case validation, and general advice and support.
- Staff at the South Western Regional Cancer Registry for help with data collection.
- Professor Sir Richard Doll, Dr Sarah Darby (ICRF, Oxford), Professor Martin Gardner (MRC Environmental Epidemiology Unit, Southampton University) and Professor John Colley (Department of Community Medicine, University of Bristol) for comments on an early draft.
- Sarah Beckhelling and Alison Hancock for assistance in data preparation and analysis.
- David Moulder (Somerset County Council) for providing population figures.
- Staff in the Department of Community Medicine for helpful comments throughout the study.
- Amanda Bowen for typing the report.

REFERENCES

1. Bowie C. (1984) Leukaemia in the West of Somerset. Somerset Health Authority. November 1984.

2. Bowie C. and Ewings P. (1987). Leukaemia in the West of Somerset: A second report. Somerset Health Authority. January 1987.

3. Glass A. G., Hill J. A. and Miller R. W. (1968). Significance of Leukaemia Clusters. The Journal of Paediatrics. 73, 1, pp 101-107.

4. Investigation of the Possible Increased Incidence of Cancer in West Cumbria. Report of the Independent Advisory Group. Chairman: Sir Douglas Black. HMSO (1984).

5. Roman E., Beral V., Carpenter L., Watson A., Barton C., Ryder, H. and Aston D. C. (1987) Childhood Leukaemia in the West Berkshire and Basingstoke and North Hampshire District Health Authorities in relation to nuclear establishments in the vicinity. British Medical Journal, 294, pp 597-602.

6. Cook-Mozaffari P. J., Ashwood F. L., Vincent T., Foreman D. and Alderson M. Cancer incidence and mortality in the vicinity of nuclear installations, England and Wales, 1950-1980. (Studies on Medical and Population Subjects, No.51). HMSO, London 1987.

7. Barclay R. (1987) Childhood Leukaemia in Wessex. Community Medicine 9, 3, pp 279-285.

8. Heasman M. A., Kemp I. W., Urquhart J. D. and Black R. (1986). Childhood Leukaemia in Northern Scotland. The Lancet, February 1st, 1986. p.266.

9. COMARE (2nd Report). Investigation of the possible increased incidence of leukaemia in young people near the Dounreay Nuclear Establishment, Caithness, Scotland. Chairman: Professor M. Bobrow. HMSO, London, 1988.

10. Shimizu Y., Kato H., Schull W. J., Preston D. L., Fujita S. and Pierce D. A. Life span study report 11 (RERF TR - 12.87) Hiroshima: Radiation Effects Research Foundation. 1987.

11. Preston D. L. and Pierce D. A. The effect of changes in dosimetry on cancer mortality risk estimates in the atomic bomb survivors. (RERF TR - 9.87). Hiroshima: Radiation Effects Research Foundation. 1987.

12. Muirhead C. R. and Darby S. C. (1987). Modelling the Relative and Absolute Risks of Radiation-Induced Cancers. Journal of Royal Statistical Society. A, 150, 2 pp 83-118.

13. Yoshimoto Y., Kato H. and Schull W. J. Risk of Cancer among children exposed in utero to A-bomb radiation, 1950-1984. The Lancet, Sept. 17, 1988. pp 665-669.

14. Bowie C. (1987). The Validity of a Cancer Register in Leukaemia Epidemology. Community Medicine. 9:1. pp 152-159.

15. Swerdlow A. J. (1986). Cancer Registration in England and Wales: Some Aspects Relevant to Interpretation of the Data. Journal of Royal Statistical Society. A, 149, 2, pp 146-160.

16. Breslow N.E. and Day N.E. (1987). Statistical Methods in Cancer Research, Vol II - The Design and Analysis of Cohort Studies (IARC Scientific Publications No. 82). Lyon, International Agency for Reseach on Cancer.

17. Kinlen L. (1988). Evidence For an Infective Cause of Childhood Leukaemia: Comparison of a Scottish New Town with Nuclear Reprocessing Sites in Britain. The Lancet, December 10, 1988, pp 1323-1327.