Oncogene (2001) 20, 4621–4628

www.nature.com/onc

Relative frequency and morphology of cancers in carriers of germline TP53 mutations

Jillian M Birch*,[1], Robert D Alston[1], Richard JQ McNally[1], D Gareth R Evans[2], Anna M Kelsey[1], Martin Harris[3], Osborn B Eden[3] and Jennifer M Varley[3]

[1]CRC Paediatric and Familial Cancer Research Group and Department of Pathology, Royal Manchester Children's Hospital, Stancliffe, Hospital Road, Manchester M27 4HA, UK; [2]Department of Medical Genetics, St. Mary's Hospital, Hathersage Road, Manchester M13 OJH, UK; [3]Department of Pathology, Academic Unit of Paediatric Oncology and CRC Cancer Genetics Group, Paterson Institute for Cancer Research, Christie Hospital NHS Trust, Wilmslow Road, Manchester M20 4BX, UK

The spectrum and frequency of cancers associated with germline TP53 mutations are uncertain. To address this issue a cohort of individuals from 28 families with Li-Fraumeni syndrome, segregating germline TP53 mutations was established. Predicted cancers were estimated by applying age, morphology, site and sex-specific UK cancer statistics to person-years at risk. Observed and predicted cancers were compared and two-sided *P*-values calculated. Cancer types occurring to excess and showing *P*-values <0.02, were designated strongly associated with germline TP53 mutations. These were removed from the data and a second round of analyses performed. Cancer types with *P*-values <0.02 and $0.02-0.05$ in the second round analyses were considered moderately and weakly associated respectively. Strongly associated cancers were: breast carcinoma, soft tissue sarcomas, osteosarcoma, brain tumours, adrenocortical carcinoma, Wilms' tumour and phyllodes tumour. Carcinoma of pancreas was moderately associated. Leukaemia and neuroblastoma were weakly associated. Other common carcinomas including lung, colon, bladder, prostate, cervix and ovary did not occur to excess. Although breast carcinoma and sarcomas were numerically most frequent, the greatest increases relative to general population rates were in adrenocortical carcinoma and phyllodes tumour. We conclude that germline TP53 mutations do not simply increase general cancer risk. There are tissue-specific effects. *Oncogene* (2001) **20,** 4621–4628.

Keywords: germline TP53 mutations; cancer frequency; cancer predisposition

Introduction

Li-Fraumeni Syndrome (LFS) is characterized by a high incidence of a variety of cancers diagnosed at young ages. On the basis of 24 families selected according to standard criteria Li *et al.* (1988) defined the main components of LFS as sarcomas, breast cancer, brain tumours, adrenocortical carcinoma and leukaemia. Results of population-based studies of families of children with sarcoma indicated that Wilms' tumour, melanoma and gonadal germ cell tumours might also be syndrome components (Hartley *et al.*, 1987, 1989, 1993). Additionally, observations in individual families led to suggestions that cancers of lung, larynx, prostate, pancreas and lymphoma might be involved (Strong *et al.*, 1987; Lynch *et al.*, 1990).

Malkin *et al.* (1990) and Srivastava *et al.* (1990) demonstrated constitutional mutations to the TP53 gene in six LFS families. Subsequent studies showed that some families with LFS did not carry TP53 mutations (Birch *et al.*, 1994; Evans *et al.*, 1998). It has also become clear that patients and families with patterns of cancer outside the strict LFS criteria may carry germline TP53 mutations (Varley *et al.*, 1997; Birch *et al.*, 1998). The question of which cancers occur with germline TP53 mutations and the frequency of these relative to incidence in the general population therefore arises.

Three previous analyses have relied on frequencies of cancers in published families and individual patients with germline TP53 mutations (Wang *et al.*, 1996; Kleihues *et al.*, 1997; Nichols *et al.*, 2001). This approach does not allow for methods of ascertainment of families, completeness or otherwise of information provided on family members and whether or not mutations have been shown to segregate with cancers in the families. These factors influence both the observed and expected patterns of cancers. In addition, time periods over which diagnoses were made, were not considered and data originating from several different countries were amalgamated. Therefore, variations in cancer incidence over time, and between populations were not taken into account. Furthermore, proband cancers were included and this also distorts results.

*Correspondence: JM Birch, CRC Paediatric and Familial Cancer Research Group, Royal Manchester Children's Hospital, Stancliffe, Hospital Road, Manchester M27 4HA, UK;
E-mail: Jillian.m.birch@man.ac.uk
Received 11 December 2000; revised 8 May 2001; accepted 9 May 2001

We have sought to overcome these difficulties by analysing cancer occurrence in 28 fully documented families segregating germline TP53 mutations which were ascertained according to pre-defined criteria for LFS (Li *et al*., 1988) or Li-Fraumeni-like (LFL) syndrome (Birch *et al*., 1994).

Results

There were 20 LFS and eight LFL families in the series including 16 with missense mutations, six with splicing mutations, three with point mutations creating a stop codon and three with deletion/insertion mutations. Four families were identified during the 1980's. The remainder have been identified since 1990. Of these, nine have not been reported previously. Proband cancers and support 1 cancers in LFL families (see Materials and methods) which by definition were of specified types, were excluded from the analyses. Five reported cancers were also excluded due to lack of verifiable information. The family cohort included 501 individuals among whom 148 eligible cancers had occurred. Histopathological material for special diagnostic review was obtained for 80 cases and pathology or other hospital records for 53. The remaining 15 cases were confirmed from cancer registrations or death certificates. Thirty-one diagnostic groups were defined, a priori, principally according to morphology. (Table 1). The total expected cancers in this cohort for ages 0–74 years based on age, sex, calender period and diagnostic group specific cancer incidence rates for England and Wales is 14.039. The overall crude relative risk is therefore 10.54. Families were selected to include at least two early onset cancers other than the proband cancer and therefore direct comparisons of observed and crude expected numbers of cancers would be inappropriate. Therefore predicted numbers of cancers were estimated for each separate diagnostic group assuming a uniform relative risk across all diagnostic groups. The predicted number (PN) is the expected number (E) in each diagnostic group multiplied by the relative risk for all cancers combined.

The distribution of observed, expected and predicted numbers of cancers within the cohort by diagnostic group is shown in Table 1. The two-sided P values are based on comparisons of observed and predicted cancers for each diagnostic group. To allow for the number of comparisons made a P-value of <0.02 was regarded as indicating a significant association. The groups osteosarcoma and chondrosarcoma, soft tissue sarcomas, central nervous system tumours, carcinoma of female breast, adrenocortical carcinoma, Wilms' tumour and malignant phyllodes tumour, which accounted for 63.5% of the total cancers, were therefore regarded as strongly associated with germline TP53 mutations. In two groups, carcinoma of lung and carcinoma of colon, the observed numbers of cancers were significantly less than predicted.

Strongly associated cancers would greatly influence the overall crude relative risk and consequently the predicted number of cancers in other groups. Therefore, to identify cancer groups more moderately or weakly associated with germline TP53 mutations, the strongly associated cancers were removed from the data. The overall relative risk, the predicted numbers of cancers in the remaining groups, and two-sided P values were then recalculated based on the reduced number of observed and expected cancers (Table 2). Carcinoma of pancreas had a P value of 0.007 in the second round of comparisons and was regarded as moderately associated. Leukaemia and peripheral nervous system tumours had P values of around 0.02. There was therefore evidence of a weak association with these groups. Carcinoma of stomach had a P value of 0.082 but all other groups had P values greater than 0.1.

Table 3 presents the relative frequencies of cancer types showing strong, moderate or weak associations with germline TP53 mutations in relation to their population expected frequencies. For each diagnostic group, relative frequency was calculated as the per cent of the total observed divided by the per cent of the total expected. Relative frequency values were calculated for each 15 year age group to determine incidence patterns by age of diagnosis. Across all ages (0–74 years) carcinoma of the female breast was numerically the most frequent cancer (25.7% of total tumours) but the relative frequency was only 1.46, reflecting the high incidence of this cancer in the general population. Soft tissue sarcoma, (12.8% of total tumours) and osteo/chondrosarcoma (6.8% of total tumours) had relative frequencies of 6.61, and 14.81 respectively, reflecting their low incidence in the general population. The most striking values were observed for two extremely rare cancers, adrenocortical carcinoma and malignant phyllodes tumour which comprised 4.7% and 1.4% of the total cancers but had relative frequencies of 77.85 and 78.06 respectively. Relative frequencies of component cancers showed considerable variation with age of diagnosis. Thus, carcinoma of the breast had the highest relative frequency (5.96) at 15–29 years, whereas for soft tissue and bone sarcomas the highest relative frequencies (6.35 and 21.54 respectively) were seen at 45–59 years.

The overall observed distribution of cancers and the overall predicted distribution were compared using simulation. Ten thousand simulations of the distribution of the 148 cancers at all ages and each 15 year age group were performed assuming that the predicted distributions were the true distributions and using the same total number of cancers. For each simulation a test statistic (TS) was calculated. An observed test statistic based on the actual data was also calculated. Figure 1 shows the distribution of TS for all ages with the value of the observed TS indicated. The probability (P) that the observed and predicted distributions are the same is the proportion of simulated values of TS that are at least as large as the observed TS. For the age range 0–74 years $P<0.0001$; for ages 0–14 years $P=0.0008$; 15–29 years $P=0.013$; 30–44 years $P=0.004$; 45–59 years $P=0.002$; and 60–74 years

Table 1 Observed and predicted cancers, ages 0–74 among 28 families with germline TP53 mutations

Diagnostic group	*Observed cancers* (O)	*Expected*[a] *cancers* (E)	*Predicted*[b] *cancers* (PN)	*Probability* (p)[c]
Carcinoma female breast	38	2.463	25.96	**0.0126**[d]
Carcinoma lung	10	2.403	25.34	0.0004[e]
Carcinoma colon	2	0.882	9.29	0.0096[e]
Carcinoma stomach	7	0.666	7.02	NS
Tumours of brain and spinal cord	14	0.639	6.74	**0.0090**[d]
Carcinoma cervix	2	0.615	6.49	0.0705
Carcinoma rectum	1	0.589	6.21	0.0230
Carcinoma bladder	1	0.550	5.80	0.0330
Leukaemia	7	0.542	5.72	NS
Carcinoma ovary	1	0.466	4.91	NS
Malignant melanoma	2	0.360	3.80	NS
Carcinoma pancreas	6	0.333	3.52	NS
Carcinoma prostate	1	0.316	3.33	NS
Carcinoma uterus	0	0.316	3.33	0.0854
Hodgkin's disease	0	0.301	3.18	0.817
Non Hodgkin's Lymphoma	2	0.397	4.19	NS
Soft tissue sarcoma	19	0.273	2.88	$<$**0.0001**[d]
Carcinoma lip, oral cavity and pharynx	1	0.251	2.65	NS
Carcinoma oesophagus	3	0.237	2.50	NS
Carcinoma kidney	3	0.237	2.50	NS
Gonadal germ cell tumours	1	0.223	2.35	NS
Multiple myeloma	1	0.155	1.64	NS
Carcinoma larynx	0	0.142	1.50	NS
Osteosarcoma and chondrosarcoma	10	0.051	0.55	$<$**0.0001**[d]
Peripheral nervous system tumours	2	0.043	0.46	0.0768
Wilms' tumour	4	0.034	0.36	**0.0005**[d]
Ewing's tumour and other peripheral PNET	0	0.016	0.17	NS
Adreno-cortical carcinoma	7	0.009	0.09	$<$**0.0001**[d]
Hepatoblastoma	0	0.004	0.04	NS
Malignant phyllodes tumour	2	0.002	0.03	**0.0003**[d]
Other and unspecified tumours	1[f]	0.485	5.13	0.0681
Totals	148	14.039	148.00	

[a]Expected number (E) of cancers based on national cancer statistics for England and Wales. [b]Predicted number (PN) of cancers assuming a uniform relative risk across all diagnostic groups,

$$\mathrm{PN} = \mathrm{E} \times \frac{\text{total observed}}{\text{total expected}}$$

[c]Probability (p) of observing (O) cancers given that the true number of cancers is PN. [d]Cancers strongly associated with mutations. [e]O Significantly less than PN. [f]one case of mesothelioma

$P = 0.28$. The pattern of cancers was therefore highly significantly different from that expected based on national cancer incidence rates allowing for age, sex and calendar period of diagnosis across all ages and in each age group except 60–74 years.

Discussion

The aims of the present study were to establish major and minor component cancers of the syndrome associated with germline TP53 mutations and to assess the level of increase in frequency of these relative to population rates. Of equal importance was also to identify cancers which did not occur at increased frequency and which are therefore not components of the syndrome.

There are now approximately 180 published examples of germline TP53 mutations in families or individual patients. However, the above aims cannot be achieved on the basis of these published reports for two main reasons. First, families and patients included in the published series have been ascertained by a variety of means and according to varied criteria, so that the frequencies of specific types of cancers are likely to be biased. Thus, published reports include studies of the incidence of germline TP53 mutations among series of patients with tumours typical of LFS thereby increasing the frequency of these cancers overall. Such studies include brain tumours (Kyritis *et al*., 1994; Chen *et al*., 1995; Zhou *et al*., 1999) breast cancer (Børresen *et al*., 1992; Sidransky *et al*., 1992; Prosser *et al*., 1992) sarcomas (Toguchida *et al*., 1992; McIntyre *et al*., 1994; Diller *et al*., 1995; Ayan *et al*., 1997) and adrenocortical carcinoma (Wagner *et al*., 1994; Sameshima *et al*., 1992; Varley *et al*., 1999). Some patients and families were selected for analysis of germline TP53 status on the basis of clusters of cancers typical of LFS e.g. familial breast cancer (Huusko *et al*., 1999; Jolly *et al*., 1994) familial brain tumours

Table 2 Observed and predicted cancers with strongly associated cancers removed from the analysis

Diagnostic group	*Observed cancers* (O)	*Expected*[a] *cancers* (E)	*Predicted*[b] *cancers* (PN)	*Probability* (p)[c]
Carcinoma lung	10	2.403	12.30	NS
Carcinoma colon	2	0.882	4.52	NS
Carcinoma stomach	7	0.666	3.41	0.0822
Carcinoma cervix	2	0.615	3.16	NS
Carcinoma rectum	1	0.589	3.02	NS
Carcinoma bladder	1	0.550	2.82	NS
Leukaemia	7	0.542	2.78	**0.0203**[d]
Carcinoma ovary	1	0.466	2.39	NS
Malignant melanoma	2	0.360	1.85	NS
Carcinoma pancreas	6	0.333	1.71	**0.0071**[e]
Carcinoma prostate	1	0.316	1.62	NS
Carcinoma uterus	0	0.316	1.62	NS
Hodgkin's disease	0	0.301	1.55	NS
Non Hodgkin's Lymphoma	2	0.397	2.03	NS
Carcinoma lip, oral cavity and pharynx	1	0.251	1.29	NS
Carcinoma oesophagus	3	0.237	1.22	NS
Carcinoma kidney	3	0.237	1.22	NS
Gonadal germ cell tumours	1	0.223	1.14	NS
Multiple myeloma	1	0.155	0.80	NS
Carcinoma larynx	0	0.142	0.73	NS
Peripheral nervous system tumours	2	0.043	0.22	**0.0209**[d]
Ewing's tumour and other peripheral PNET	0	0.016	0.08	NS
Hepatoblastoma	0	0.004	0.02	NS
Other and unspecified tumours	1	0.485	2.49	NS
Total cancers	54	10.555	54.00	

[a]Expected number (E) of cancers based on national cancer statistics for England and Wales. [b]Predicted number (PN) of cancers assuming a uniform relative risk across all diagnostic groups,

$$\mathrm{PN} = \mathrm{E} \times \frac{\text{total observed}}{\text{total expected}}$$

[c]Probability (p) of observing (O) cancers given that the true number of cancers is PN. [d]Cancers weakly associated with mutations. [e]Cancers strongly associated with mutations

Table 3 Relative frequencies of cancers showing strong, moderate or weak association with germline TP53 mutations by age

	Age groups (years)																	
	0–14			*15–29*			*30–44*			*45–59*			*60–74*			*0–74*		
Diagnostic group	% *Total* O	% *Total* E	*Rel. freq.*	% *Total* O	% *Total* E	*Rel. freq.*	% *Total* O	% *Total* E	*Rel. freq.*	% *Total* O	% *Total* E	*Rel. freq.*	% *Total* O	% *Total* E	*Rel. freq.*	% *Total* O	% *Total* E	*Rel. freq.*
Carcinoma female breast	0	0.1	–	34.5	5.7	**5.96**	38.0	29.0	**1.31**	24.2	21.2	**1.14**	14.3	12.5	**1.14**	25.7	17.5	**1.46**
Tumours of brain and spinal cord	13.8	23.5	0.59	10.3	10.6	0.97	12.0	5.6	**2.14**	3.0	3.5	0.88	0	2.0	–	9.5	4.6	**2.08**
Leukaemia	10.3	34.4	0.31	6.9	8.8	0.78	2.0	2.9	0.68	0	1.9	–	14.3	2.1	**6.89**	4.7	3.9	**1.22**
Carcinoma pancreas	0	0.05	–	0	0.3	–	4.0	1.1	**3.51**	9.1	2.5	**3.70**	14.3	3.5	**4.09**	4.1	2.4	**1.71**
Soft tissue sarcoma	20.7	6.2	**3.32**	6.9	6.1	**1.11**	16.0	2.8	**5.72**	9.1	1.4	**6.35**	0	0.8	–	12.8	1.9	**6.61**
Sarcomas of bone and cartilage	10.3	3.1	**3.30**	17.2	2.9	**5.95**	2.0	0.3	**5.88**	3.0	0.1	**21.54**	0	0.1	–	6.8	0.5	**14.81**
Peripheral nervous tumours	6.9	6.2	**1.12**	0	0.42	–	0	0.07	–	0	0.04	–	0	0.02	–	1.4	0.3	**4.39**
Wilms' tumour	13.8	5.9	**2.35**	0	0.09	–	0	0.006	–	0	0.002	–	0	0.002	–	2.7	0.2	**11.01**
Adreno-cortical carcinoma	20.7	0.3	**75.84**	0	0.1	–	0	0.08	–	3.0	0.05	**60.36**	0	0.03	–	4.7	0.06	**77.85**
Malignant phyllodes tumour	0	<0.001	–	0	0.02	–	2.0	0.04	**54.95**	3.0	0.02	**149.25**	0	0.007	–	1.4	0.01	**78.06**

Relative frequency (Rel. freq.) $= \frac{\%\ \text{Total O}}{\%\ \text{Total E}}$ (for definitions see Table 1). Relative frequency >1 indicates that crude relative risk is higher than relative risk for all cancers in respective age group

(Vital *et al*., 1998; Lubbe *et al*., 1995) and multiple primary tumours (Russo *et al*., 1994; Malkin *et al*., 1997). Second, in many reports of germline TP53 mutations in families, full family histories, including the number and sex of unaffected family members, current ages or ages at death of unaffected as well as affected individuals have not been given (Frebourg *et al*., 1995; Rines *et al*., 1998; Goi *et al*., 1997).

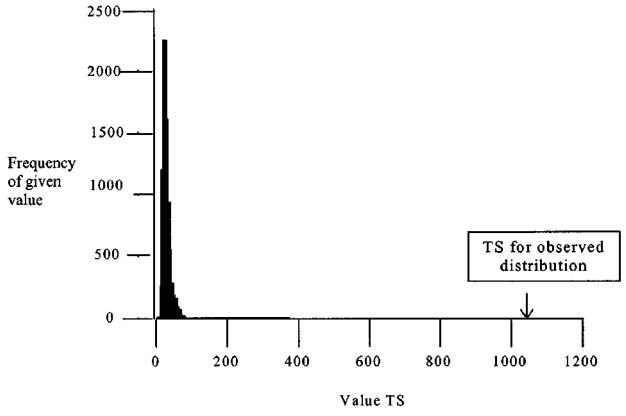

Figure 1 Histogram of predicted test statistic (TS) for distribution of cancers, ages 0–74 years

Therefore, the expected frequency of cancers cannot be estimated.

In the present study, these difficulties were addressed by including only families ascertained according to standard criteria and by excluding proband cancers in all families and the first support cancer in LFL families. By definition these must be sarcoma, breast cancer, brain tumour, leukaemia or adrenocortical carcinoma. All other cancers in the families included in this study were completely unselected for cancer type. The data are therefore not biased towards excess occurrence of certain cancers. However, all families were selected for TP53 analysis because of high cancer incidence but the methodology used in the present study was devised to overcome this problem.

There was an unusually high degree of diagnostic confirmation and histopathological material was available for review in over half the cases. Virtually complete information was available on all eligible family members, including sex, dates of birth, death and diagnosis. Multiple family members have been tested for germline TP53 status and the patient cohort comprised a defined set of relatives. Thus, each individual included in the analysis was of known mutation status or a close relative. We therefore have an exceptionally well-documented, standardized cohort of families segregating fully characterized germline TP53 mutations. In addition, and uniquely, detailed national individual level cancer registration statistics, which specify morphological types of cancers were available to us. We were therefore able to examine cancers by morphological type rather than primary site, allowing for age, sex and calendar period of occurrence.

On the basis of all these detailed and accurate data, we were able to demonstrate unequivocally that the distribution of cancers within our cohort was highly significantly different from the expected cancer distribution. This was true of cancers across all ages and in each age group below 60 years. Therefore, we can conclude that the effect of a germline TP53 mutation is not simply to increase cancer risk in general. This represents the first formal demonstration that tissue specific effects on cancer risks are present in such families.

The analysis identified seven cancer types as being strongly associated with germline TP53 mutations. These seven include five of the six principal cancer types identified by Li *et al*. (1988) as being components of LFS (carcinoma of female breast, tumours of brain and spinal cord, soft tissue sarcoma, osteosarcoma but not Ewing's tumour, and adrenocortical carcinoma). However, leukaemia did not emerge as a major component cancer. Wilms' tumour was also identified as strongly associated with such mutations and it is interesting to note that all four cases arose in families with mutations affecting splicing. Six of the 28 families (21%) carried splicing mutations (Varley *et al*., 2001). Such mutations are under-represented among published families and patients since in many studies intron-exon boundaries were not included in gene sequence analyses (Frebourg *et al*., 1995; Verselis *et al*., 2000). This may account for the paucity of Wilms' tumour cases in published families.

A novel finding was the association with malignant phyllodes tumour. We selected this rare malignant breast tumour for analysis since in LFS families breast cancers are frequent and stromal tissue prone to malignant change. In the UK cancer registration data, malignant phyllodes tumour comprised 0.2% of all malignant breast tumours. This is consistent with a reported incidence of 2.1 per million women in a population-based study from the USA (Bernstein *et al*., 1993) Although our observation is based on only two cases, there are two other cases in the literature (Mazoyer *et al*., 1994) plus a further case described as stromal sarcoma of the breast (Goi *et al*., 1997). Furthermore, in one of our families, there was a case of benign phyllodes tumour which was not included in the analysis.

Carcinoma of the pancreas emerged as being moderately associated. This is the only common adult onset cancer in addition to breast carcinoma which appears to be a component of the TP53 syndrome. There was no association with the majority of common carcinomas including lung, colon, rectum, cervix, bladder, ovary and prostate, but there was very weak evidence of excess cases of stomach carcinoma. However, the possibility of increased risk of these cancers in carriers of germline TP53 mutations who are also at risk as a result of exposure to environmental carcinogens (e.g. tobacco) or other contributing genetic factors cannot be excluded. The small number of such cancers occurring at unusually early ages in our series and in the literature, may have arisen as a result of combined genetic and environmental factors.

Peripheral nervous system tumours, based on two cases of neuroblastoma in young children showed a weak association. There are reports of three other families with germline TP53 mutations in the literature

which include a case of neuroblastoma (Brugières *et al*., 1993; Pivnick *et al*., 1998; Porter *et al*., 1992) and there is a further case in our series in a family which did not fulfill the eligibility criteria for the present study. It therefore seems likely that neuroblastoma is a rare component tumour. Gonadal germ cell tumours and melanoma did not emerge as component neoplasms of the TP53 syndrome in spite of evidence suggesting excess frequency in clinically defined LFS families.

Cancers included in the cohort are unselected with respect to morphology and primary site. The cohort therefore allows an unbiased assessment of the pattern of cancers associated with germline TP53 mutations relative to that expected in a population with similar age and sex structure. The results shown in Table 3 indicate that in individuals with germline TP53 mutations, carcinoma of the female breast, soft tissue sarcoma, tumours of brain and spinal cord and osteo/chondrosarcoma are the most important cancers in numerical terms especially at young ages. However, in terms of relative risk adrenocortical carcinoma and malignant phyllodes tumour appear to be the most strongly associated with such mutations. In addition, our data indicate a very clear trend of increasing risk with decreasing age at onset. The ratios of observed cancers to the crude expected number of cancers in the age groups 0–14 years, 15–29 years, 30–44 years, 45–59 years and 60–74 years are 51.2, 32.7, 20.2, 6.7 and 1.4 respectively (P value for trend < 0.0005).

In summary, in this study of 28 extensively documented families with germline TP53 mutations we have demonstrated that carcinoma of female breast, tumours of brain and spinal cord, soft tissue sarcoma, osteosarcoma and chondrosarcoma and adrenocortical carcinoma are strongly associated with the mutations. In addition to these anticipated principal component cancers, we also identified Wilms' tumour and malignant phyllodes tumour as syndrome components. Carcinoma of the pancreas showed a moderate association. Leukaemia and neuroblastoma were weakly associated. There was no evidence that other common carcinomas, melanoma and gonadal germ cell tumours represented component cancers. Therefore, cancers with the highest frequencies of somatic mutations to TP53, e.g. carcinomas of colon and lung, are rarely seen in carriers of germline mutations. Reasons for this are not clear but may reflect the timing of events and relationship to exogenous carcinogens. Thus in lung and colorectal carcinoma, TP53 mutation is a late event in the carcinogenic process but in tumours developing in carriers of germline mutations it is axiomatic that the TP53 mutation is the first event. We conclude that the TP53 germline mutation syndrome consists of a specific spectrum of mainly uncommon cancers.

These data provide guidance in formulating management policies in patients with germline TP53 mutations and in selecting suitable families for analysis of germline TP53 mutation status prior to enrolment in predictive testing programmes.

Materials and methods

Families were considered eligible for inclusion if they carried a fully characterized germline TP53 mutation and fulfilled the criteria for either classic LFS as follows: an individual with bone or soft tissue sarcoma diagnosed before 45 years of age (designated the proband), a first degree relative with any cancer diagnosed before 45 years (designated support case 1) and another first or second degree relative in the same lineage with any cancer in this age range (designated support case 2) (Li *et al*., 1988); or Li-Fraumeni-like (LFL) syndrome, previously defined by us as: a proband with childhood cancer or sarcoma, brain tumour or adrenocortical carcinoma under 45 years of age with one first or second degree relative with a typical LFS cancer (sarcoma, breast cancer, brain tumour, leukaemia or adrenocortical carcinoma) at any age (support case 1) and another first or second degree relative in the lineage with any cancer below 60 years of age (support case 2 Birch *et al*., 1994). There were 28 eligible families ascertained from three sources: (i) Studies of cancer incidence in the families of children with bone or soft tissue sarcoma, included in the Manchester Children's Tumour Registry (eight families). (ii) The United Kingdom Children's Cancer Study Group (UKCCSG) patient register. The UKCCSG comprises a network of paediatric oncology centres which collectively treat more than 80% of all incident cases of cancer in children in the UK. The register includes information on neoplasms in parents and siblings. (Ten families). (iii) Families with multiple cases of cancer attending regional genetics centres in the UK. Families or individuals requesting genetic counselling with respect to cancer risk would be referred to one of these centres. (Ten families). Families originally identified as multi-case breast cancer families were excluded. Research on these families was carried out with their consent and with the approval of the local research ethics committee.

A cohort was established which included the probands, their first and second degree relatives and descendants in the cancer lineage in each family. For each individual the following details were obtained as applicable: sex, dates of birth, neoplasms with dates of diagnosis, dates of death and causes of death. Methods for defining the cohort are given in detail elsewhere (Birch *et al*., 1998). Reported cases of neoplasms were verified by obtaining copies of histopathological reports, medical records, death certificates and cancer registrations. In addition, histopathological material was obtained for diagnostic review and to provide a source of DNA. Peripheral blood samples were obtained from as many family members as possible. Coding and non-coding regions of the TP53 gene were sequenced in constitutional samples using automated methods. The molecular methods are described in detail elsewhere (Varley *et al*., 1997).

Cancers which were eligible for inclusion in the analysis were all malignant primary tumours (excluding non-melanoma skin cancer) and any central nervous system tumour. All proband cancers and support case 1 cancers in LFL families were excluded. For determination of the expected numbers of cancers, person years at risk for members of the cohort were calculated up to 75th birthday or death. Anonymized individual cancer registration records were supplied on CD ROM by the Office for National Statistics (London) for new cases of cancer in England and Wales diagnosed from 1971 until 1992 inclusive. Cancer diagnoses in the national data and in the patient cohort were allocated topography and morphology codes, according to Manual of Tumor Nomenclature & Coding (American Cancer Society Inc. 1968) for cases diagnosed from 1971–1978 and International Classifi-

cation of Diseases for Oncology edition 1 (World Health Organisation, 1976) from 1979. A classification scheme for grouping the cancer incidence data by diagnosis was devised primarily based on morphology, but also using topography to sub-divide carcinomas by primary site. Intracranial germ cell tumours were classified with central nervous system tumours.

Expected numbers of cancers within the cohort for each diagnostic group were calculated by applying the national statistics by 5 year age group from 0–74 years, sex and 10 year calendar period to the person-years at risk. The cancer groups were then ranked according to frequency of expected cancers. Groups for further analysis were selected as follows: the main LFS component cancers defined by Li *et al*. (1988) (osteosarcoma and chondrosarcoma, soft tissue sarcomas, brain tumours, breast carcinoma, adrenocortical carcinoma and leukaemia); other cancers suggested to be components of LFS on the basis of systematic studies (Wilms' tumour, melanoma and gonadal germ cell tumours); cancers suggested to be components on the basis of *ad hoc* observations on families with LFS (carcinomas of lung, larynx, prostate and pancreas and lymphoma); cancers which for a priori reasons might be expected to occur with the syndrome (other paediatric tumours and malignant phyllodes tumour); the 10 cancer groups with the highest expected numbers in addition to the above. This gives a total of 30 specified cancer groups, plus all other and unspecified cancers (Table 1). Details of the morphology and topography code specifications for these groupings may be obtained from the authors.

Predicted numbers of cancers (PN) were then estimated for each separate diagnostic group assuming a uniform relative risk across all diagnostic groups within the families. PN is then the expected number in each group (E_1, E_2 etc) multiplied by the relative risk for all cancers combined.

To identify specific cancer types associated with TP53 germline mutations, the following procedure was adopted. For each pre-specified cancer group individually, the observed number (O) was compared with the predicted number (PN). Assuming that the total number of cancers is fixed (n), then numbers in individual cancer group will have a binomial distribution. The probability of a case being from the pre-specified cancer group is PN/n. The statistical significance of the observed number (O) of cancers is then assessed by calculating a two-sided probability. Diagnostic groups with excess numbers of cases and with *P*-values of less than 0.02 were regarded as showing a strong association with TP53 germline mutations. A significance level of less than 0.02 was selected since there were 31 groups in the analysis and less than one significant result at this level would be expected by chance.

The overall distribution of cancers among the diagnostic groups and the distribution predicted given the same relative risk for every cancer type were compared using simulation. Ten thousand simulations of the distribution of cancers were performed using the same total numbers of cancers and assuming the true distribution was that predicted. For each simulation a test statistic,

$$TS = \sum \frac{(O - PN)^2}{PN}$$

was calculated. Asymptotically the values of TS follow a chi-squared distribution. An observed test statistic based on the actual data was then calculated.

The probability (P) that the overall observed and predicted distributions are the same was calculated as the proportion of simulated values of TS that are at least as large as the observed value. This procedure was repeated for distributions of cancers in the age groups 0–14 years, 15–29 years, 30–44 years, 45–60 years and 60–74 years.

Acknowledgments

This research was funded by grants from the Cancer Research Campaign. Jillian M Birch is Cancer Research Campaign Professorial Research Fellow in Paediatric Oncology and Osborn B Eden is Cancer Research Campaign Professor of Paediatric Oncology at the University of Manchester. We thank the families who took part in the study and the medical personnel who provided clinical details. We are particularly grateful to the many pathologists who supplied material for analysis and to the Office for National Statistics in London for cancer registration data.

References

Ayan I, Luca JW, Jaffe N, Strong LC and Hansen MF. (1997). *Oncol. Rep.,* **4,** 679–683.

Bernstein L, Deapen D and Ross RK. (1993). *Cancer,* **71,** 3020–3023.

Birch JM, Blair V, Kelsey AM, Evans DG, Harris M, Tricker KJ and Varley JM. (1998). *Oncogene,* **17,** 1061–1068.

Birch JM, Hartley AL, Tricker KJ, Prosser J, Condie A, Kelsey AM, Harris M, Morris Jones PH, Binchy A, Crowther D, Craft AW, Eden OB, Evans DGR, Thompson E, Mann JR, Martin J, Mitchell ELD and Santibanez-Koref MF. (1994). *Cancer Res.,* **54,** 1298–1304.

Birch JM, Heighway J, Teare MD, Kelsey AM, Hartley AL, Tricker KJ, Crowther D, Lane DP and Santibanez-Koref MF. (1994). *Br. J. Cancer,* **70,** 1176–1181.

Bol FJ, Sleddens HF and Dinjens WN. (1998). *Diag. Mol. Pathol.,* **7,** 295–301.

Børresen A-L, Anderson TI, Garber J, Barbier-Piraux N, Thorlacius S and Eyfjörd J. (1992). *Cancer Res.,* **52,** 3234–3236.

Brugières L, Gardes M, Moutou C, Chompret A, Meresse V, Martin A, Poisson N, Flamant F, Bonaïti-Pellie C, Lemerle J and Feunteun J. (1993). *Cancer Res.,* **53,** 452–455.

Chen P, Iavarone A, Fick J, Edwards M, Prados M and Israel MA. (1995). *Cancer Genet. Cytogenet,* **82,** 106–115.

Diller L, Sexsmith E, Gottlieb A, Li FP and Malkin D. (1995). *J. Clin. Invest,* **95,** 1606–1611.

Evans SC, Mims B, McMasters KM, Foster CJ, de Andrade M, Amos CI, Strong LC and Lozano G. (1998). *Hum. Genet.,* **102,** 681–686.

Frebourg T, Barbier N, Yan Y, Garber JE, Dreyfus M, Fraumeni J, Li FP and Friend SH. (1995). *Am. J. Hum. Genet.,* **56,** 608–615.

Goi K, Takagi M, Iwata S, Delia D, Asada M, Donghi R, Tsunematsu Y, Nakazawa S, Yamamoto H, Yokota J, Tamura K, Saeki Y, Utsonomiya J, Takahashi T, Ueda R, Ishioka C, Eguchi M, Kamata N and Mizutani S. (1997). *Cancer Res.,* **57,** 1895–1902.

Hartley AL, Birch JM, Kelsey AM, Marsden HB, Harris M and Teare MD. (1989). *Cancer Genet. Cytogenet,* **42,** 221–226.

Hartley AL, Birch JM, Marsden HB and Harris M. (1987). *J. Med. Genet.,* **24,** 664–668.

Hartley AL, Birch JM, Tricker K, Wallace SA, Kelsey AM, Harris M and Morris Jones PH. (1993). *Cancer Genet. Cytogent,* **67,** 133–135.

Huusko P, Castren K, Launonen V, Soini Y, Paakkonen K, Leisti J, Vahakangas K and Winqvist R. (1999). *Cancer Genet. Cytogenet,* **112,** 9–14.

Jolly KW, Malkin D, Douglass EC, Brown TF, Sinclair AE and Look AT. (1994). *Oncogene,* **9,** 97–102.

Kleihues P, Schäuble B, Hausen zur A, Estève J and Ohgaki H. (1997). *Am. J. Pathol,* **150,** 1–13.

Kyritis AP, Bondy ML, Xiao M, Berman EL, Cunningham JE, Lee PS, Levin VA and Saya H. (1994). *JNCI.,* **86,** 344–349.

Li FP, Fraumeni Jr. JF, Mulvihill JJ, Blattner WA, Dreyfus MG, Tucker MA and Miller RW. (1988). *Cancer Res.,* **48,** 5358–5362.

Li Y-J, Sanson M, Hoang-Xuan K, Delattre J-Y, Poisson M, Thomas G and Hamelin R. (1995). *Int. J. Cancer,* **64,** 383–387.

Lubbe J, von Ammon K, Watanabe K, Hegi ME and Kleihues P. (1995). *Brain Path.,* **5,** 15–23.

Lynch HT, Radford B and Lynch JF. (1990). *Oncology,* **47,** 75–79.

Malkin D, Friend SH, Li FP and Strong LC. (1997). *New Engl. J. Med.,* **336,** 734.

Malkin D, Li FP, Strong LC, Fraumeni JF Jr, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA and Friend SH. (1990). *Science,* **250,** 1233–1238.

Mazoyer S, Lalle P, Moyret-Lalle C, Marcais C, Schraub S, Frappaz D, Sobol H and Ozturk M. (1994). *Oncogene,* **9,** 1237–1239.

McIntyre JF, Smith-Sorensen B, Friend SH, Kassell J, Borresen AL, Yan YX, Russo C, Sato J, Barbier N, Miser J, Malkin D and Gebhardt MC. (1994). *J. Clin. Oncol.,* **12,** 925–930.

Nichols KE, Malkin D, Garber JE, Fraumeni Jr JF and Li FP. (2001). *Cancer Epi. Biom. Prev.,* **10,** 83–87.

Pivnick EK, Furman WL, Velagaleti GVN, Jenkins JJ, Chase NA and Ribiero RC. (1998). *J. Med. Genet.,* **35,** 328–332.

Porter DE, Holden ST, Steel CM, Cohen BB, Wallace MR and Reid R. (1992). *J. Bone. Joint Surg.,* **74,** 883–886.

Prosser J, Porter D, Coles C, Condie A, Thompson AM, Chetty U, Steel CM and Evans HJ. (1992). *Brit. J. Cancer,* **65,** 527–528.

Rines RD, van Orsouw J, Sigalas I, Li FP, Eng C and Vijg J. (1998). *Carcinogenesis,* **19,** 979–984.

Russo CL, McIntyre J, Goorin AM, Link MP, Gebhardt MC and Friend SH. (1994). *Med. Pediat. Oncol.,* **23,** 354–358.

Sameshima Y, Tsunematsu Y, Watanabe S, Tsukamoto T, Kawa-ha K, Hirata Y, Mizoguchi H, Sugimura T, Terada M and Yokota J. (1992). *JNCI.,* **84,** 703–707.

Sidransky D, Tokino T, Helzlsouer K, Zehnbauer B, Rausch G, Shelton B, Prestigiacoma L, Vogelstein B and Davidson N. (1992). *Cancer Res.,* **52,** 2984–2986.

Srivastava S, Zou Z, Pirollo K, Blattner W and Chang EH. (1990). *Nature,* **348,** 747–749.

Strong LC, Stine M and Norsted TL. (1987). *JNCI.,* **79,** 1213–1220.

Toguchida J, Yamaguchi T, Dayton SH, Beauchamp RL, Herrera GE, Ishizaki K, Yamamuro T, Meyers PA, Little JB, Sasaki MS, Weichselbaum RR and Yandell DW. (1992). *New Eng. J. Med.,* **326,** 1301–1308.

Varley JM, Evans DGR and Birch JM. (1997). *Br. J. Cancer,* **76,** 1–14.

Varley JM, McGown G, Thorncroft M, James LA, Margison GP, Forster G, Evans DGR, Harris M, Kelsey AM and Birch JM. (1999). *Am. J. Hum. Genet,* **65,** 995–1006.

Varley JM, McGown G, Thorncroft M, Santibàñez-Koref MF, Kelsey AM, Tricker KJ, Evans DGR and Birch JM. (1997). *Cancer Res.,* **57,** 3245–3252.

Varley JM, Attwooll C, White G, McGown G, Thorncroft M, Kelsey AM, Greaves M, Boyle J and Birch JM. (2001). *Oncogene,* **20,** 2647–2654.

Verselis SJ, Rheinwald JG, Fraumeni Jr JF and Li FP. (2000). *Oncogene,* **19,** 4230–4235.

Vital A, Bringuier P-P. Huang H, Galli FS, Rivel J, Ansoborlo S, Cazauran J-M, Taillandier L, Kleihues P and Ohgaki H. (1998). *J. Neuropathol Exp. Neurol,* **57,** 1061–1069.

Wang Q, Lasset C, Sobol H and Ozturk M. (1996). *Int. J. Cancer,* **65,** 554–557.

Wagner J, Portwine C, Rabin K, Leclerc J-M, Narod SA and Malkin D. (1994). *JNCI.,* **86,** 1707–1710.

Zhou X-P, Sanson M, Hoang-Xuan K, Robin E, Taillander L, Cornu P, Delattre J-Y, Thomas G and Hamelin R. (1999). *Ann. Neurol,* **46,** 913–916.