Joint Effects of Germ-Line p53 Mutation and Sex on Cancer Risk in Li-Fraumeni Syndrome

Chih-Chieh Wu,¹ Sanjay Shete,¹ Christopher I. Amos,¹ and Louise C. Strong²

¹Department of Epidemiology and ²Section of Clinical Cancer Genetics, Department of Molecular Genetics, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

Abstract
Germ-line p53 mutations have been identified in most families with Li-Fraumeni syndrome (LFS). For germ-line p53 mutation carriers, there is considerable variability with respect to age of cancer onset and tumor type, suggesting that additional genetic effects influence the clinical severity and tumor spectrum. To identify factors that might contribute to the observed heterogeneity in time to onset, we used segregation analysis to analyze the joint effects of germ-line p53 mutations and risk modifier(s) on cancer incidence. We studied 159 kindreds, ascertained through probands who had been diagnosed with childhood soft-tissue sarcoma before 16 years of age, survived >3 years after diagnosis, and treated at The University of Texas M.D. Anderson Cancer Center (Houston, TX) from 1944 to 1975. This unique cohort has been followed systematically for >20 years and has had germ-line p53 mutation testing in probands and extended family members. The analyses revealed that germ-line p53 mutations and sex had significant effects on cancer risk: men with p53 mutations had 151-fold higher odds of developing cancer than did those without mutations [95% confidence interval (95% CI), 60-380], and women with p53 mutations had 1,075-fold higher odds than did men with mutations (95% CI, 2.5-20.3). These findings provide quantitative cancer risk assessments for LFS families. (Cancer Res 2006; 66(16): 8287-92)

Introduction
Li-Fraumeni syndrome (LFS) is a rare familial cancer syndrome characterized by a high frequency of early-onset and diverse tumor types and an increased frequency of multiple primary tumors (1, 2). The initial description of the syndrome included the aggregation in families of soft-tissue sarcomas (STS), osteosarcomas, female breast cancer, brain tumors, leukemias, and adrenocortical carcinomas (3–7). Further studies have also shown greatly increased risks for many other common cancers, including prostate (7) and lung cancer (8). In addition, individuals from families with LFS have an increased propensity to develop multiple primary tumors (3).

Mutations in p53 are the most common tumor-specific genetic alteration in human neoplasms; they have been identified in >50% of human cancers, including many different cancer types. Germ-line mutations of p53 in chromosome 17p13.1 have been identified in 50% to 70% of families with LFS (9–11). As some LFS patients and their relatives affected with invasive cancers do not have germ-line mutations in p53, it is unlikely that mutations at a single locus, p53, can account for all clinical phenotypes of LFS. Even in the presence of a stable germ-line p53 mutation segregating in a family, the age of cancer onset and multiplicity of tumors within a family is variable, suggesting the presence of additional risk modifier(s). A small minority of LFS families with no p53 mutations has been noted to have mutations in CHEK2 (12). In addition, linkage to a region of chromosome 1q has been noted for some larger families with LFS who did not have CHEK2 or p53 mutations (13). Given these observations, we proposed to determine the quantitative effects of p53 mutations on cancer risk and to evaluate the effects of modifier(s) of cancer risk.

To characterize and evaluate cancer risk related to germ-line p53 mutations, we analyzed the cancer incidence in 159 extended pedigrees ascertained through 3-year survivors of childhood STS (3, 7, 14). We have collected systematically cancer incidence data for these families for >20 years. We analyzed samples from 107 families for p53 mutations and identified 7 with germ-line mutations. In those kindreds, a total of 63 germ-line p53 mutation carriers have been identified (7, 9, 15). A particular advantage of this study is that a very large percentage of family members have been genotyped without consideration of affection status, thus yielding more precise estimates of cancer risk associated with p53 mutations.

Given the diversity of tumor types observed in excess in LFS (6, 7), we included all invasive cancers, except nonmelanoma skin cancer and in situ carcinoma as a single combined phenotype. Assuming that cancer incidence follows a logistic distribution, we did segregation analysis allowing for genetic and nongenetic covariates, such as germ-line p53 mutation, sex, generation cohorts, and relationship cohorts. Our goals were to assess the associations between cancer incidence and several possible genetic and nongenetic covariates, to evaluate the cancer risk attributable to significant covariates in a quantitative way, and to provide the 95% confidence intervals (95% CI) associated with risk estimates.

Unlike most family studies, in which p53 mutations were genotyped only in patients meeting the clinical definition of LFS (16), the probands in this study were selected because of the childhood sarcoma and not because of any familial cancer. In kindreds with a p53 germ-line mutation, genotyping was extended to the relevant first-degree relatives of mutation carriers without regard to affection status. These families have been followed prospectively now for >20 years. Therefore, our analysis could well characterize and evaluate the cancer risk related to germ-line p53 mutations for familial cancer studies. These results will be valuable in determining the genetic etiology in the presence or absence of germ-line p53 mutations, localizing or identifying non-p53 disease susceptibility loci or risk-modifying loci, and clinical counseling.
Materials and Methods

Study population. The study population used for this investigation consisted of 159 patients with STS, who have been diagnosed before 16 years of age, survived $>$3 years after diagnosis, and treated at The University of Texas M. D. Anderson Cancer Center (Houston, TX) from 1944 to 1975, and their extended relatives, trimmed to include grandparents, aunts and uncles, parents, full siblings, and offspring of all probands. Kindreds in which a p53 germ-line mutation was identified were extended to include first-degree relatives of all individuals with p53 germ-line mutations. Because extension through mutation status was not done with respect to the phenotype, this approach to extending the family should not introduce an ascertainment bias during the segregation analysis.

For this analysis, we included all invasive cancers, except nonmelanoma skin cancers and in situ carcinoma. These disease criteria are broader than the classic LFS component tumors but are based on observations of diverse cancers occurring in excess in p53 germ-line mutation carriers (6–8). All cancers included in the analysis were confirmed by medical records or death certificates. Persons were considered at risk from date of birth to date of cancer diagnosis, death, lost to follow-up, study termination (December 31, 2001, the study termination date for medical record and death certificate documentation), or age of 75 years, whichever came first. The evaluation of cancer incidence was truncated at age of 75 years because of the limited reliability of cancer rates at older ages.

The final data set for the 159 kindreds consisted of 3,034 individuals, with 257 men and 215 women affected. This total is composed of 283 individuals from the 7 extended kindreds with a p53 germ-line mutation (7), including 63 mutation carriers (11 with no cancer), 1,813 individuals from the remaining 100 kindreds for which the probands tested negative for p53, and 938 individuals from the 52 kindreds for which no sample was available for testing. Because germ-line p53 mutations are rare in general population and because the distributions of affected relatives in the 52 families are very close to those in 100 families that are p53 negative, we considered a total of 132 families as p53-negative families in our analyses. The number of males and females in each group was similar.

All probands (or "surrogates," who were affected close relatives in cases in which the proband was deceased) were tested for p53. We then tested adult first-degree relatives of p53-positive probands. Next, the p53 mutation kindreds were extended to test at-risk adult relatives of p53 mutation carriers. In kindreds in which the proband tested negative for the p53 mutation, no other family members were specifically tested for p53 by full sequencing, but for higher risk kindreds, p53 was ruled out by linkage analysis (13). More details of the methods of data collection, cancer incidence overall, the identification of germ-line p53 mutations, and the frequencies of site-specific cancers for those with and without mutations have been described elsewhere (3, 7, 8, 14, 17).

Statistical methods. On the assumption that conditional on genotypes age of onset of first cancer follows a logistic distribution, we did maximum likelihood segregation analyses to assess the significance of several possible cancer risk factors, including p53 mutations, sex, family generation, and family relationship. Next, we quantitatively evaluated cancer risk attributable to significant covariates and obtained the associated 95% CIs by inverting the Fisher information matrix, which was obtained as part of the maximum likelihood estimation of variables.

The analysis was conducted using REGTL, a module of the Statistical Analysis for Genetic Epidemiology (SAGE) software release 3.1 (18). Under model 1 of the program, class A regressive models were used that assumed that the genotypes influence age of onset of the phenotype through the location and scale variables of the logistic distribution but do not influence susceptibility (19). The only form of major gene inheritance allowed for in the program was a single locus with two alleles.

The time intervals for the regressive models were the time between birth and diagnosis of the first cancer for case subjects and the time between birth and death, last contact, age of 75 years, or study termination date of December 31, 2001, whichever came first, for those without cancer. Single-ascertainment correction was used in the analysis because the families were identified through patients with STS (20).

The variables of regressive logistic models for genetic analysis used include $q_A$, the frequency of the putative high-risk allele A; $\tau_u$, the transmission probability of allele A for genotype $u$; $x$, the age adjustment coefficient; and $\beta_u$, the baseline variable for genotype $u$. The scale and location variables of the logistic distribution are $x$ and $\beta_u$, respectively, where $\beta_u$ is the natural logarithm of the odds of being affected versus being unaffected by genotype $u$ when other components and covariates are zero. Under the assumption of Mendelian inheritance, the values of $\tau$ are fixed at $\tau_{AA} = 1.0$, $\tau_{AB} = 0.5$, and $\tau_{BB} = 0.0$. The effects of genotype AA are the same as those for genotype AB, as reflected by the baseline variables $\beta_{AA} = \beta_{AB}$ for the Mendelian dominant model; $\beta_{BB} = \beta_{AB}$ for the Mendelian recessive model; and $\beta_{BB} = 0.5(\beta_{AA} + \beta_{AB})$ for the Mendelian additive model. The effects of genotypes AA, AB, and BB decrease in that order, as reflected by $\beta_{AA} \geq \beta_{AB} \geq \beta_{BB}$ for the Mendelian decreasing (major-gene) model. The Mendelian arbitrary (major-gene) model does not put restrictions on the values of $\beta$. No-major-gene (sporadic) models assume that baseline risk is not influenced by genotype but can vary among populations.

The genetic and nongenetic covariates considered in the analysis were described as follows:

(a) Germ-line p53 mutation: those with mutations were coded as 1 and those with wild-type (WT) as 0. In 7 extended kindreds with a germ-line p53 mutation, 63 were carriers, 138 were WT, and 82 were at risk for being a mutation carrier but had unknown genotypes. Those with unknown genotypes in p53-positive families were given the values of probabilities of carrying a p53 mutation, which were calculated using Bayes rule to incorporate the mutation statuses of any relatives who were tested for mutations within the families.

(b) Relationship cohort: the probands were coded as 0. Their first-, second-, and third-degree and higher degree relatives and marry-ins were coded as 1, 2, and 3, respectively.

(c) Generation cohort: the generation of probands was coded as 0. Their children and grandchildren were coded as 1 and 2, respectively. Their parents were coded as $-1$, and their grandparents and earlier ancestors were coded as $-2$.

(d) Sex: males were coded as 0 and females as 1.

Hypothesis testing. The logarithm ln(L) of the maximum likelihood of the data was computed for each model. The likelihood ratio test (LRT) was used to test a specific model against the baseline model, which is usually the general model, in which the values of $\tau$ used are arbitrary to identify the best fit to the data for the general model. The specific model serves as the null model and the baseline model as the alternative model. The LRT was computed as follows: $\text{LRT} = -2\left(\ln(L_{\text{maximun}}) - \ln(L_{\text{baseline}})\right)$, where LRT approximately follows a $\chi^2$ distribution with degrees of freedom equal to the difference in the numbers of independent variables estimated in the two models. Another method of model comparisons is Akaike's information criteria (AIC), defined as $\text{AIC} = -2 \ln(L_{\text{maximun}}) + 2$ (number of independent variable estimated). The model with the lowest AIC value and fewer estimated variables is generally considered the most parsimonious one. The LRT was also used to test for significance of the covariates included in the models where the model that included additional covariate(s) is regarded as the baseline model. Because some variables (such as some of the transmission variables) are fixed at boundaries for some hypothesis tests, the significance tests that we have applied yield conservative P values.

The genetic models for the modifier(s) of cancer risk underlying LFS remain unknown. Thus, it would not be appropriate to use highly restricted models for them, such as the Mendelian dominant, additive, or recessive models, when testing for the significance of covariates. We chose to test for the significance of covariates on the basis of the Mendelian decreasing model. The best fit of the Mendelian decreasing model suggests the presence of a single gene with reduced penetrance, which could be used to underlie multifactorial etiology (e.g., polygene or oligogene effects) accounting for many familial cancers (21). The decreasing model could also represent the composite effects of several genes or environmental factors that are strongly correlated among family members.
Results

We first assessed the significance of germ-line p53 mutations in our study. Letting the model that includes p53 mutations serve as the null model and the model with no covariates serves as the model with its environmental model, for which the values of $s$ and $b$, which were defined previously. The results of the minimum model and null model for LRT testing are presented in columns 3 and 2 of Table 1, respectively.

Given the observations of Chompret et al (22) and Hwang et al (7) about increased cancer risk in female p53 mutation carriers, we wanted to investigate the effects of sex on cancer risk in p53 germ-line mutation carriers. We used the covariate p53*sex, the product of p53 and sex, which estimates the excess of cancer risk in female carriers over male carriers (note that sex is coded as 1 for females and 0 for males in our scheme). Letting the sex-specific p53 effect model be the usual p53 effect model, for which the values of $s$ and $b$, which were defined previously. The null hypothesis of the usual p53 effect model was rejected, suggesting that the sex difference on cancer risk is highly significant in people with germ-line p53 mutations differentially than the general population. This analysis shows that the sex-specific p53 effect model provides a better fit to the data.

Next, we incorporated the covariate of sex alone in the models with or without p53 mutations, but the results did not improve the fit of the model. To investigate the possible phenomenon of genetic anticipation or a birth cohort effect for mutation carriers, as suggested by some studies of LFS (23, 24), we included the covariates p53*generation (the product of p53 and generation effect) and generation in the analysis. According to the data shown in columns 6 and 7 of Table 1, the comparisons with the sex-specific p53 model in column 4 using the LRT, p53*generation and generation were not significant factors in the models.

The use of the relationship cohort was based on our previous observation of familial cancer aggregations in first-degree relatives (3). Including the relationship covariate in the model did not significantly improve the likelihood value of the model as shown by comparing the two models in columns 4 and 8 of Table 1.

After we identified p53 and p53*sex to be the only two significant covariates in our segregation analysis, we further compared the sex-specific p53 model with its general model, for which the transmission probabilities $\tau$ are arbitrary (as opposed to the values of $s$ and $b$). Letting the sex-specific p53 effect model be the null model, as shown in columns 4 and 3 of Table 1, respectively, the $\chi^2$ value for LRT with 1 degree of freedom was 6.53, giving $P = 0.011$. The null hypothesis of the usual p53 effect model was rejected, suggesting that the sex difference on cancer risk is highly significant in people with germ-line p53 mutations differentially than the general population. This analysis shows that the sex-specific p53 effect model provides a better fit to the data.

<table>
<thead>
<tr>
<th>Variable Covariates included in the models</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_a$ 0.0003 0.0984 0.0157 0.0270 0.0154 0.0156 0.0152</td>
</tr>
<tr>
<td>$\tau_{AA}$ 1.0 1.0 1.0 1.0 1.0 1.0 1.0</td>
</tr>
<tr>
<td>$\tau_{AB}$ 0.5 0.5 0.5 0.103 0.5 0.5 0.5</td>
</tr>
<tr>
<td>$\tau_{BB}$ 0.0 0.0 0.0 0.024 0.0 0.0 0.0</td>
</tr>
<tr>
<td>$\beta_{AA}$, female $-2.80$ $-4.95$ $-1.92$ $-4.92$ $-1.91$ $-1.93$ $-1.58$</td>
</tr>
<tr>
<td>$\beta_{AB}$, female $-3.30$ $-8.52$ $-6.87$ $-6.03$ $-6.90$ $-6.89$ $-6.55$</td>
</tr>
<tr>
<td>$\beta_{BB}$, female $-8.41$ $-8.98$ $-9.03$ $-10.66$ $-9.06$ $-9.05$ $-8.76$</td>
</tr>
<tr>
<td>$\beta_{AA}$, male $-2.37$ $-7.48$ $-4.63$ $-5.08$ $-4.63$ $-4.63$ $-4.43$</td>
</tr>
<tr>
<td>$\beta_{AB}$, male $-4.35$ $-7.48$ $-7.12$ $-7.66$ $-7.15$ $-7.14$ $-6.93$</td>
</tr>
<tr>
<td>$\beta_{BB}$, male $-8.25$ $-9.53$ $-8.71$ $-10.01$ $-8.74$ $-8.73$ $-8.42$</td>
</tr>
<tr>
<td>$\sigma$ 0.0898 0.0938 0.0937 0.1116 0.0942 0.0934 0.0937</td>
</tr>
<tr>
<td>p53 $-$ 5.53 4.56 5.02 4.69 4.58 4.63</td>
</tr>
<tr>
<td>p53sex $-$ $-$ 1.44 1.96 1.34 1.46 1.46</td>
</tr>
<tr>
<td>Relationship $-$ $-$ $-$ $-$ $-$ $-0.14$</td>
</tr>
<tr>
<td>p53*generation $-$ $-$ $-$ $-$ $-$ $-0.20$</td>
</tr>
<tr>
<td>Generation $-$ $-$ $-$ $-$ $-$ $-0.31$</td>
</tr>
<tr>
<td>$-2 \ln(L)$ 3.97679 3.85470 3.84817 3.83873 3.84711 3.84809 3.84726</td>
</tr>
<tr>
<td>Degrees of freedom $-$ 1 $-$ 1 $-$ 1 $-$ 1</td>
</tr>
<tr>
<td>$\chi^2$ $-$ 122.09 6.53 9.44 1.06 0.08 0.91</td>
</tr>
<tr>
<td>$P$ $-$ &lt;0.0001 0.011 0.024 0.303 0.777 0.340</td>
</tr>
</tbody>
</table>

NOTE: $\tau_{AA} = 1.0$, $\tau_{AB} = 0.5$, and $\tau_{BB} = 0.0$ are fixed, except the general model on column 5.
fixed at one single value (i.e., \( \tau_{AA} = \tau_{AB} = \tau_{BB} \)). The \( \chi^2 \) value for the LRT was 7.33 that gives \( P = 0.026 \) with 2 degrees of freedom (data not shown), suggesting that the environmental model was rejected at a 0.05 nominal significance level. We also compared the general sex-specific model with its sporadic model. The \( \chi^2 \) value for the LRT was 27.99, which gives \( P = 0.0005 \) with 8 degrees of freedom (data not shown), suggesting that the sporadic model was rejected at a 0.01 nominal significance level. In conclusion, the general sex-specific model was our most plausible model, and the AIC criterion confirmed this result as shown in Table 1. Our plausible model showed the value of \( \tau_{AA} \) was bounded at 1.0, \( \tau_{AB} \) had a reduced transmission probability of 0.103 (compared with 0.5 expected under a Mendelian model), and \( \tau_{BB} \) was very close to 0, suggesting that modifier(s) of cancer risk other than germ-line p53 mutations may influence familial cancer incidence.

The estimated covariate coefficients for the general sex-specific p53 effect model revealed that the p53 and p53*sex covariates were very strongly associated with cancer incidence. The coefficient of the p53 covariate was 5.02, indicating that the odds ratio (OR) for being affected was \( e^{\beta_{p53}} = 151.41 \) for men with germ-line p53 mutations. The coefficient of p53*sex was 1.96, indicating that the OR for being affected was \( e^{\beta_{p53} \cdot 1.96} = 1074.92 \) for women with p53 mutations. Women with p53 mutations had 7.1 (\( e^{1.96} \))-fold higher odds on cancer incidence than did men with p53 mutations. Consequently, the significance of the p53*sex covariate with coefficient 1.96 in regressive logistic models shows that women with p53 mutations were substantially younger at first cancer diagnosis than were men, and we will show this result in details later. The 95% CIs of the ORs for cancer in people with p53 mutations were 60 to 380 and 358 to 3,229 for men and women, respectively. The 95% CI of the OR for the excess of cancer in women versus men with p53 mutations was 2.5 to 2.03. One can note that the coefficients reflecting effects of p53 mutations and sex change little among the various models that have been fitted that allowed for residual familial correlation.

It is noteworthy that, in the models, including the p53 effect as a covariate, the effect of p53 mutations on cancer incidence is assessed by the magnitude of its coefficient, and the values of \( a_i, \tau_u, \alpha, \) and \( \beta_u \) describe the generic model of disease transmission for modifier(s) of cancer risk contributing to cancer incidence as shown in the last five columns of Table 1.

According to the data in Table 1, we can calculate the cumulative cancer risk by p53 mutation, sex, genotype, and age using the expression (19, 25, 26)

\[
\gamma \left(1 \div \exp(-\alpha \times a_i - \beta_u - \zeta x_1 - \ldots - \zeta x_v)\right)
\]

where \( \alpha \) and \( \beta_u \) represent the scale and location variables of the logistic distribution, respectively; \( a_i \) is the age at examination, both for unaffected \( i \)th individual and for affected \( i \)th individual for whom the age of onset is unknown; \( \gamma \) is the susceptibility (the probability that a randomly selected individual from the population will ever become affected as age goes to infinity); and \( \xi \) represents the coefficient of covariate \( x \). The values of baseline (location) variable \( \beta \) for our most plausible model shown in column 5 of Table 1 indicated that the estimated age of onset for heterozygotes was substantially younger for women than for men but not for homozygotes. This trend was clear when we compared the cumulative cancer risk for genotype AB between women and men as shown in Table 2.

The models presented in Table 1 were based on the assumption of sex dependence for genotype-specific baseline variable \( \beta \). We analyzed the same data set, assuming sex independence for \( \beta \), and found that each of the five models shown in the last five columns of Table 1 was better than the corresponding sex-independent model (data not shown). The associations between cancer incidence and the covariates, p53 or p53*sex, remained strong, and these were the only two significant covariates in the sex-independent models. These results could provide insight into the genetic basis of LFS and could be useful in designs of localizing non-p53 susceptibility genes or risk modifiers.

We also tested for genetic interaction between p53 mutations and unobserved modifier(s) of p53 for the 7 families with germ-line p53 mutations using SAGE 5.0 (27). Based on the Mendelian arbitrary model, the values of \((-2)P\) likelihood were 796.25 for the general sex-specific p53 model with genotypic p53 mutation interaction and 802.87 for the corresponding model without genetic interaction. The \( \chi^2 \) value for LRT with 2 degrees of freedom was 6.62, which gives \( P = 0.037 \). This is modest evidence of an interaction between germ-line p53 mutations and p53 modifier for the seven families with p53 mutations. The variable coefficients of p53 and p53*sex, which are centered at mean 0.0, were 17.66 and 1.92, respectively. The variable coefficients of genotype-dependent effects of p53 were 0.08, −0.21, and 0.12, respectively. Accordingly, p53 and p53*sex had very strong effects, decreasing the ages of onset by 87 and 12 years, respectively. In contrast, the genotype-dependent effects of p53 in terms of age-of-onset changes were −0.41, 1.05, and −0.64 for p53_AA, p53_AB, and p53_BB, respectively.

Although we found significant genetic interaction in the models at a 0.05 nominal significance level, the genotype-dependent fluctuations of p53 were very small compared with the main effects of p53 and p53*sex. The assessment of cancer risk attributable to significant covariates, which is assumed to be genotype independent, were essentially unchanged, whether or not we allowed for p53 interactions with a modifier locus. Therefore, the results we report in Table 1 are accurate. Furthermore, inclusion of the interaction effects led to flatness in the likelihood surface, which suggests that at least some of the variables in the segregation model became confounded with each other, so this extended model may not yield reliable results.

**Discussion**

We have a unique study population that was not ascertained because of characteristics suggestive of LFS (other than a proband with a childhood STS), and we have followed-up systematically the cohort for >20 years. Therefore, our analysis could well characterize and evaluate the cancer risk related to germ-line p53 mutations. Consequently, we applied a segregation analysis that allowed us to associate the simultaneous effects of germ-line p53 mutations and unobserved modifier(s) of cancer risk underlying LFS with cancer incidence in these families, possibly accounting for genetic modification of cancer risk among or within these families. This design jointly determines the most likely genetic model of modifier(s) of cancer risk contributing to LFS and the effects of p53 mutations and other genetic or nongenetic factors on cancer risk. Most clinical observations and case reports structured to identify germ-line p53 mutations determined the excess cancer risk in patients with mutations or their relatives using standardized incidence ratios. In contrast, we quantitatively evaluated the effects...
of germ-line p53 mutations and other factors on cancer risk using OLs in a maximum likelihood segregation analysis, which was particularly designed for family data.

We identified a significantly higher cancer risk in people with germ-line p53 mutations in our models than in those without mutations. Individuals with germ-line p53 mutations had 151-fold higher odds of developing cancer than those with no mutations in men and 1,075-fold higher odds in women. More importantly, we found a significant sex difference in cancer risk among those with germ-line p53 mutations; women with mutations had 7.10-fold higher odds of having cancer than did men with mutations. These results indicate that those with mutations were younger at first cancer diagnosis than were those with no mutations and that age of onset at first cancer risk for women with mutations was significantly lower than in men as shown in Table 2.

In addition to the analysis done on the basis of the Mendelian decreasing model, as shown in Table 1, we also did this analysis using the Mendelian arbitrary models (the values of $\beta$ are arbitrary as opposed to the constraint $\beta_{AA} \geq \beta_{AB} \geq \beta_{BB}$ for the Mendelian decreasing model) and found similar results to the ones shown in Table 1. The only two significant covariates were p53 mutations and p53sex. Based on the Mendelian arbitrary model, the values of (−2)$\chi_2$ likelihood were 3.838.05 for the general sex-specific p53 model and 3.848.02 for the sex-specific p53 model (data not shown). The models provided only negligible improvement in likelihood values over those based on the Mendelian decreasing model. The $\chi^2$ value for LRT with 3 degrees of freedom was 9.977 that gives $P = 0.019$, suggesting that the sex-specific p53 model was rejected at a 0.05 nominal significance level. The general sex-specific p53 model was statistically better at a 0.05 nominal significance level.

In contrast to the Mendelian decreasing model, which could underlie the composite effects of genetic or environmental factors, the Mendelian arbitrary model could describe possible genetic heterogeneity, allowing for situations, such as $\beta_{AA} \leq \beta_{AB} \leq \beta_{BB}$ (overdominance). It is possible that the heterozygotes have a more extreme phenotype than the disease-allele homozygotes, but strong support or evidence of overdominance is needed to be believable as due to a major gene effect. This hypothesis did not seem to apply in our case. The Mendelian decreasing model provided a better fit to our data.

We conclude that the general sex-specific p53 effect model was the most plausible model in our study on the basis of the following reasons: the general sex-specific p53 model was the most plausible model either using the Mendelian decreasing model or Mendelian arbitrary model, and when comparing the sex-specific p53 model with its general model, the $\chi^2$ test for LRT allows 3 degrees of freedom to accommodate the arbitrary values of the three genotype-dependent transmission probabilities $\tau$. However, the values of $\tau$ are bounded within the interval [0, 1], and the actual degrees of freedom are <3. As a result, the $\chi^2$ test for LRT with 3 degrees of freedom tends to be conservative in this case (the actual significance level is lower than the nominal significance level). Thus, the $P = 0.024$ and 0.019 for comparing the sex-specific p53 model with its general model using the Mendelian decreasing model and Mendelian arbitrary model, respectively, may have been too conservative. The general sex-specific p53 model seems to provide a significantly better fit than the sex-specific p53 model either using the Mendelian decreasing or Mendelian arbitrary model.

Furthermore, the most plausible models that are the general sex-specific p53 model in our analysis, assuming sex dependence for the baseline variable $\beta$, were significantly better in likelihood than the corresponding sex-independent models at a 0.01 nominal significance level using either the Mendelian arbitrary or decreasing model (data not shown). These results suggest that additional modifier(s) of cancer risk, which was responsible for the familial cancer incidence, was more likely to be sex specific.

Some analytic approaches have suggested an anticipation or birth cohort effect in LFS (23, 24). Given that most of the offspring generations of our probands were <21 years, most had not undergone genetic testing, so we were limited in our ability to discriminate the effects of a p53 mutation carrier. Thus, our findings for the family relationship, generation, and p53 generation covariates should not be considered conclusive. Additional years of observation and genotyping are needed to confirm these findings and determine those effects.

We chose to model the age-specific risks depending on genotype and not the susceptibility variables based on consideration of biological processes causing an increased risk for cancer. According to either the two-stage or multistage models of cancer, genetic susceptibility for cancer predisposition affects the rate at which cancers develop not the baseline susceptibility. In addition, existing statistical treatments of other cancers using REGTL (21, 28) have typically shown that modeling genetic susceptibility as influencing the time to onset for cancer provides a better fit to data for familial cancers than modeling genetic susceptibility through baseline probabilities.

It is noteworthy that, given that all invasive cancers except nonmelanoma skin cancer and in situ carcinoma were included as a single combined phenotype in our analysis, we actually assumed that all cancer cases at 17 different sites were at equal risk in the setting of our models. The frequencies of site-specific incidence have been described in our previous work (7). The association between lung cancer and smoking-related cancer and cigarette

### Table 2. Age-specific cumulative cancer risk by p53 mutation, sex, and genotype

<table>
<thead>
<tr>
<th>Mutation/sex</th>
<th>Ages/genotype</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>AB</td>
<td>BB</td>
</tr>
<tr>
<td>p53+</td>
<td>&lt;20</td>
<td>0.898</td>
<td>0.399</td>
</tr>
<tr>
<td></td>
<td>&lt;40</td>
<td>0.988</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td>&lt;60</td>
<td>0.999</td>
<td>0.983</td>
</tr>
<tr>
<td>p53−</td>
<td>&lt;20</td>
<td>0.055</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>&lt;40</td>
<td>0.351</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>&lt;60</td>
<td>0.834</td>
<td>0.276</td>
</tr>
</tbody>
</table>
smoking in germ-line p53 mutation carriers was analyzed and published elsewhere (8). Other assessments of cancer risk for site-specific incidence seem impractical in our case because of the sample size limitations of site-specific data.

The identification of a sex difference in cancer risk in patients with germ-line p53 mutations is important. Although a difference in cancer risk by gender was first noted by Chompret et al. (22) and later by Hwang et al. (7), this approach provides a more quantitative assessment of the difference in risk. A future goal is to understand the mechanism by which sex modulates cancer risk among individuals with germ-line p53 mutations.

Hwang et al. (7) investigated cancer risk in germ-line p53 mutation carriers using the Kaplan-Meier methods, the Cox's proportional hazards models, and the standardized incidence ratios. However, these methods only partially accommodate family structures in the analysis. The methods used by Hwang et al. adjust for intrafamilial correlations, assuming all individuals within a family have the same degree of correlation, and therefore do not allow for more complex correlation structure in the extended families we have studied, which included 56 relatives with mutations from 7 families. The maximum likelihood segregation models we used in this report are designed to analyze family data. In addition, the transmission of disease allele from one generation to another is efficiently modeled by our analysis.

Chompret et al. (22) used a different statistical approach to estimate the cancer risk in carriers of germ-line p53 mutations ascertained through childhood cancer patients. They included data from 268 index patients who were 18 years old or younger and had any type of solid malignant tumors; we ascertained families from 268 index patients who were 18 years old or younger and had another is efficiently modeled by our analysis.

We thank Dr. Guillermina Lozano for the p53 mutation testing; Xiaojun Zhou, Gloria Robertson, Doris Sembera, Mi Won Park, and Phyllis Begin for the data and sample collection; and especially the family members for their participation over the years.

Acknowledgments
Received 11/29/2005; revised 5/5/2006; accepted 6/14/2006.

Grant support: NIH grant P01 CA34936.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

References
18. S.A.G.E. Cleveland: statistical analysis for genetic epidemiology, release 3.3. Computer program package from the Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University; 1998.
27. S.A.G.E. Cleveland: statistical analysis for genetic epidemiology, release 3.3. Computer program package from the Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University; 2003.

Joint Effects of Germ-Line p53 Mutation and Sex on Cancer Risk in Li-Fraumeni Syndrome

Chih-Chieh Wu, Sanjay Shete, Christopher I. Amos, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/66/16/8287

Cited articles
This article cites 26 articles, 8 of which you can access for free at:
http://cancerres.aacrjournals.org/content/66/16/8287.full.html#ref-list-1

Citing articles
This article has been cited by 11 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/66/16/8287.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.