Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Research article:

Estimating *TP53* Mutation Carrier Probability in Families with Li-Fraumeni Syndrome Using LFSPRO

Gang Peng[1], Jasmina Bojadzieva[2], Mandy L. Ballinger[3], Jialu Li[1], Amanda L. Blackford[4], Phuong L. Mai[5], Sharon A. Savage[5], David M. Thomas[3], Louise C. Strong[2], Wenyi Wang[1]

1. Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston TX 77030
2. Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston TX 77030
3. The Kinghorn Cancer Center and Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia
4. Johns Hopkins Sidney Kimmel Comprehensive Cancer Center, Baltimore, MD 21287
5. Clinical Genetics Branch, Division of Cancer Genetic and Epidemiology, National Cancer Institute, Bethesda, MD 20892
6.

Running title: *TP53* Mutation Probability Estimation

Key words: *de novo* mutation, risk prediction, family history, Mendelian models, genetic counseling

Financial Support:

G.Peng is supported in part by the Cancer Prevention Research Institute of Texas through grant number RP130090. J.Bojadzeiva and L.C.Strong are supported in part by the U.S. National Institutes of Health through grant P01CA34936. D.M.Thomas is supported by a Senior Research Fellowship from the Australian National Health and Research Council (APP103929). The International Sarcoma Kindred Study has

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

received support from the Rainbows for Kate Foundation, the Liddy Shriver Sarcoma Initiative, the Australian National Health and Medical Research Council (APP1004017) and Cancer Australia (APP1067094). W.Wang is supported in part by the Cancer Prevention Research Institute of Texas through grant number RP130090, and by the U.S. National Cancer Institute through grant numbers 1R01CA174206, 1R01 CA183793 and P30 CA016672. S.A.Savage and P.L.Mai are supported by funding from the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, U.S. National Cancer Institute.

Correspondence to:

Wenyi Wang, PhD

Associate Professor

Department of Bioinformatics and Computational Biology

The University of Texas MD Anderson Cancer Center

Email: wwang7@mdanderson.org Phone: 713-792-5377

Website: http://odin.mdacc.tmc.edu/~wwang7

Conflict of interest

The author(s) declare that they have no conflict of interest.

Article info

Abstract word count: 250

Body text word count: 4314

Figures: 2

Tables: 4

Supplementary data: 8 figures and 5 tables

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Abstract

Background: Li-Fraumeni syndrome (LFS) is associated with germline *TP53* mutations and a very high lifetime cancer risk. Algorithms that assess a patient's risk of inherited cancer predisposition are often used in clinical counseling. The existing LFS criteria have limitations, suggesting the need for an advanced prediction tool to support clinical decision making for *TP53* mutation testing and LFS management.

Methods: Based on a Mendelian model, LFSPRO estimates *TP53* mutation probability through the Elston-Stewart algorithm, and consequently estimates future risk of cancer. With independent datasets of 1,353 tested individuals from 867 families, we evaluated the prediction performance of LFSPRO.

Results: LFSPRO accurately predicted *TP53* mutation carriers in a pediatric sarcoma cohort from MD Anderson Cancer Center (MDACC) in the US, the observed to expected ratio (OE) = 1.35 (95%CI, 0.99 to 1.80); area under the receiver operating characteristic curve (AUC) = 0.85 (0.75 to 0.93); a population-based sarcoma cohort from the International Sarcoma Kindred Study (ISKS) in Australia, OE = 1.62 (1.03 to 2.55); AUC = 0.67 (0.54 to 0.79); and the NCI LFS study cohort, OE = 1.28 (1.17 to 1.39); AUC = 0.82 (0.78 to 0.86). LFSPRO also showed higher sensitivity and specificity than the classic LFS and Chompret criteria. LFSPRO is freely available through the R packages LFSPRO and BayesMendel.

Conclusions: LFSPRO shows good performance in predicting *TP53* mutations in individuals and families in varied situations.

Impact: LFSPRO is more broadly applicable than the current clinical criteria and may improve clinical management for individuals and families with LFS.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Introduction

Li-Fraumeni syndrome (LFS) is an autosomal dominant cancer predisposition syndrome(1,2) commonly manifested as soft tissue and bone sarcomas, breast cancer, brain tumors, leukemia and adrenal cortical carcinoma(3-5). Clinical criteria were established to describe families with classic LFS(3) (Figure S1).

Germline mutations in the tumor suppressor gene *TP53* are present in about 70% of families meeting classic LFS criteria (6,7) (Figure S1). In these families, the lifetime cancer penetrance for *TP53* mutation carriers is 90%(8,9), increasing to nearly 100% for women by age 70 years, mostly because of breast cancer(10). Given this extremely high lifetime cancer risk, tools to identify who to test for germline *TP53* mutations may be useful. One study demonstrated significantly improved survival of *TP53* mutation carriers using clinical surveillance and whole-body imaging(11). While several surveillance studies are underway(12-16), whole-body imaging is recommended in US clinical management guidelines(17). Emerging evidence supports improved outcomes with targeted risk management(18).

Classic LFS criteria have high specificity for predicting *TP53* mutations, but low sensitivity. Further criteria assist in identifying potential germline *TP53* mutation carriers(7,19-22), but are relatively insensitive to *de novo* mutations. Chompret criteria (Figure S1)(20-22) have evolved to identify *de novo TP53* mutations, which may affect up to 20% of carriers(23). They have higher sensitivity than classic LFS criteria, but decreased specificity.

Widely used, classic LFS and Chompret criteria provide binary results and only apply to affected counselees and families with LFS-spectrum cancers. To address these limitations, we developed LFSPRO, which uses a Mendelian risk prediction model(24) that has been used successfully in familial breast and ovarian(25), bowel(26), and pancreatic(27) cancers and melanoma(28). We introduced *de novo* mutations into the Mendelian model to incorporate this

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

relatively frequent occurrence in LFS individuals(29), and validated our model using three independent datasets.

Materials and Methods

Model Development

LFSPRO estimates the probability of any designated family member carrying a *TP53* mutation based on a detailed family cancer history (Table S1) and values of two sets of parameters, the prevalence and penetrance for *TP53*, respectively. The prevalence of *TP53* mutations is expressed as Pr(*G*), where *G* denotes genotype, which could be wildtype (denoted as 0), a heterozygous mutation (denoted as 1) or a homozygous mutation (denoted as 2). Pr(*G*) for the three genotypes can be derived from the prevalence of mutated alleles using the Hardy-Weinberg equilibrium (HWE). The penetrance is the probability of developing cancer at a given age, for *TP53* mutation carriers and noncarriers, respectively. LFSPRO further estimates the individual's future cancer risk if currently asymptomatic, using the estimated carrier probabilities and the penetrance values.

The probability of the counselee's genotype G_0 given the family cancer history is estimated via the following formula(24).

$$\Pr(G_0 \mid H_0, H_1, \cdots, H_n) = \frac{\Pr(G_0)\Pr(H_0,\ H_1, \cdots, H_n|G_0)}{\sum_{G_0}\Pr(G_0)\Pr(H_0,\ H_1, \cdots, H_n|G_0)}$$

$$\Pr(H_0, H_1, \cdots, H_n|G_0) = \sum_{G_1\cdots G_n}\Pr(H_0, H_1, \cdots, H_n|G_0, G_1, \cdots\ G_n)\Pr(G_1, \cdots\ G_n|G_0)$$

$$= \sum_{G_1\cdots G_n}\left[\prod_{i=0}^{n}\Pr(H_i|G_i)\right]\Pr(G_1, \cdots\ G_n|G_0)$$

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Here, H_i denotes the cancer history for individual *i*, including the cancer status, diagnosed as having cancer, healthy or dead; and age, age at diagnosis, at last contact or at death. We use n to denote the total number of counselee's relatives in a family. We may consider the calculation of $\Pr(G_0|H)$ as updating the population prevalence $\Pr(G_0)$ by incorporating family cancer history. The term $\Pr(H_0,H_1, \ldots ,H_n|G_0)$ is the probability of the phenotypes for the whole pedigree given the genotype of the counselee. Because this is complex to evaluate directly, it is computed as a weighted average of the probabilities of family history given each possible genotype configuration of all relatives, where the weights are the probabilities of the genotype configuration based on Mendelian transmission. The conditional probabilities are products of the individual probability distributions of penetrance when we assume conditional independence. The posterior probability calculation is performed using the Elston-Stewart, or peeling, algorithm(30). This algorithm splits the whole pedigree into anterior and posterior parts according to the individual of interest. The anterior part relates to the parents of the individual, and the posterior part relates to the child/children of the individual. The probability of the anterior and posterior parts can be estimated recursively, such that the posterior genotype probability is calculated according to the probability of the anterior part and the posterior part. The peeling algorithm characterizes Mendelian transmission using a transmission matrix $\Pr(G_i|G_{fi},G_{mi})$, which is the probability of the genotype for individual *i*, given the genotypes of his/her father and mother. When Mendel's law is strictly followed, as in previous Mendelian models(26-28,31), $\Pr(G_i=1 \text{ or } 2|G_{fi}=0, G_{mi}=0)=0$. However, LFSPRO accounts for the occurrence of *de novo* mutations: when neither parent carries a *TP53* mutation, we still consider a transmission probability of $2a$ that their child may become a carrier of a *TP53* germline mutation, where a denotes the probability of obtaining a new mutated allele.

The net risk of developing any invasive cancer is calculated as a weighted sum of the penetrance, where weight is the estimated probability of each genotype configuration. We

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

convert these net risks to clinically relevant absolute risks by accounting for the chance of dying of other causes(32), for which we obtain estimates from the SEER 2008–2010 report(33).

Penetrance and prevalence are inputs for LFSPRO

We used previous penetrance estimates for *TP53* mutation carriers and noncarriers from six large pediatric sarcoma families at the University of Texas MD Anderson Cancer Center (MDACC), which are not included in the test cohort(10). For Pr(*G*=1,2), the prevalence of *TP53* germline mutations, we found estimates from three sources. *i*) Based on reports from two population studies(29,34), we used 0.0001 as a reasonable estimate of the general population prevalence. However, recently available commercial tests for genes that predispose to hereditary cancer syndromes have produced new data on *TP53* germline mutation prevalence. *ii*) A study by Myriad Genetics gave a frequency of 1 in 1600 (82 of 135,609) in a population (affected or unaffected with cancer) undergoing genetic testing(35). *iii*) A study by Ambry Genetics found 1 in 300 (69 of 22,226) in a cancer patient population undergoing genetic testing(36). To compare the true prevalence of *TP53* mutations in clinic populations to that in the general population, we modeled LFSPRO outputs, assuming the above prevalence estimates. Gonzalez et al.(23) estimated the proportion of *de novo* mutations within germline mutations of *TP53* as 7%-20%, with paternal or maternal origin. We used both 7% and 20% to calculate the prevalence of *de novo TP53* mutations from the three sets of prevalence data described above. Using three independent datasets (Tables 1, 2), we evaluated the performance of LFSPRO with inputs of the six combinations of prevalence estimates (Table S2). In the LFSPRO software, the user can modify the inputs of penetrance and prevalence.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Test cohorts

We tested our method using three unique datasets, none of which were used in model development (Table 1, 2). *i*) Patients with childhood soft-tissue sarcoma or osteosarcoma treated at MDACC (Houston, TX) from 1944–1983 and their extended family members comprised the 'pediatric sarcoma cohort.' Details on data collection, specific cancers and germline testing have been published(8,9,37-39). The pediatric sarcoma cohort comprised 183 unrelated families with 2,553 individuals. Eleven of the 183 kindreds had at least one individual with a confirmed germline *TP53* mutation; the rest had no confirmed mutation carriers. Due to the nature of ascertainment through probands with pediatric sarcoma, there are more cancer patients in the carrier families than in noncarrier families (average of 14 vs. 1.5, including the proband), resulting in a higher average age at diagnosis in the former set of families (Tables 1, 2). *ii*) The International Sarcoma Kindred Study (ISKS) dataset(40) is a prospective cohort of patients with adult-onset sarcoma and their extended families, recruited from six treatment centers across Australia. This 'adult sarcoma cohort' comprised 582 separate families with 16,977 individuals. Nineteen of the kindreds had at least one individual with a confirmed *TP53* mutation. *iii*) The National Cancer Institute (NCI) LFS cohort (NCT01443468) is from a long-term prospective, natural history study that started in 2011, and includes individuals meeting classic LFS or modified criteria(19), having a pathogenic germline *TP53* mutation or a first- or second-degree relative with a mutation, or a personal history of choroid plexus carcinoma, adrenocortical carcinoma, or at least 3 primary cancers. A detailed family history questionnaire was obtained, including birth date, vital status, date or age at death if deceased, history of cancer, and if positive, type and year of diagnosis or age at diagnosis, for all first-, second-, and third-degree relatives and any other extended family members for whom the information was available. There were 2,676 individuals from 102 families included in the NCI LFS cohort (see eMethods for details on the datasets). We defined case subjects as affected only if diagnosed

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

with malignant tumors, excluding non-melanoma skin cancers, but including tumors of the adrenal cortex, choroid plexus, ovary (granulosa) and breast carcinoma in situ. Cancers were classified by site: LFS spectrum (osteosarcomas, soft tissue sarcomas, breast, brain, adrenal, or lung cancers and leukemia) and non-LFS spectrum (prostate, colon, kidney, or thyroid cancers and others).

Mutation testing

Peripheral blood samples were collected after obtaining informed consent. The probands' *TP53* status was determined by PCR sequencing of exons 2-11 at MDACC(8). Additionally, high-resolution melt analysis and multiplex ligation-dependent probe amplification had been performed in the ISKS and NCI cohort data to detect large deletions or genomic rearrangements(40).

At MDACC, if a *TP53* mutation was identified, then all first-degree relatives of the proband (affected and unaffected by cancer) and any other family member at risk of carrying the familial mutation were tested. Extending germline testing based on mutation status and not on phenotype of family members should not introduce ascertainment bias during analysis(8,10). Individuals unavailable for testing (largely deceased) linked to or between confirmed mutation carriers were considered obligate mutation carriers. No other family member was tested when the proband tested negative.

Within the NCI cohort, for individuals who had testing prior to enrollment, copies of the clinical TP53 test reports were obtained and verified by the study team. For individuals actively participating in the protocol and not previously tested, clinical genetic testing was performed after enrollment. All at-risk family members of individuals who tested positive for a mutation (either prior to enrollment or on study) were offered the option of having site-specific testing through the study. No testing was offered to relatives if the proband tested negative for a TP53

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

mutation. Obligate carriers were considered mutation-positive regardless of whether testing was done.

Within the ISKS cohort, *TP53* testing was performed on probands, and on family members whoever are willingness to participate in the study. The test procedures were as described previously(40).

Evaluation of LFSPRO

We used LFSPRO to calculate the carrier probabilities of *TP53* mutations for all tested individuals (*TP53* positive and negative). We used observed:estimated ratios (OEs) to evaluate the calibration and receiver operating characteristic (ROC) curves to evaluate our model's discrimination ability. A high area under the ROC curve (AUC) indicates that we can find a point on the ROC curve for determining the *TP53* mutation carrier with a high true positive rate and low false positive rate. The OE is the ratio between the number of observed *TP53* mutation carriers and the summation of the estimated probability of *TP53* mutation carriers. The ideal situation is when the number of observed carriers equals the number estimated (OE = 1). When there is more than one family member tested, we do not include mutation information of anyone in the LFSPRO calculation for each family member. We estimated 95% confidence intervals (CIs) on our summary measures based on 10,000 bootstrap samples, using the family instead of the individual as the sample unit to account for dependence in outcomes among individuals within a family. For the same individuals, we applied classic LFS and Chompret criteria and compared binary outcomes (mutation carrier or not) with the test results. All analyses were performed using the open-source environment R (http://cran.r-project.org).

Evolution of family history in the pediatric cohort

Information on the 183 kindreds from the pediatric sarcoma cohort at MDACC had been collected over a range of 22 to 62 years (Table S3). We arranged the test data, starting when

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

the proband was first diagnosed with invasive cancer, and including only family members, their ages and their cancer status that existed during that period. We repeated our analysis at 10-year intervals based on the projected family history of that period, which resulted in 4 time points: the proband's ascertainment date was 0, then 10, 20, and the last contact date, with the length of the last time interval varying across families. These time points are relative and thus corresponded to different calendar years for each family, as illustrated in Figure S2 in a subset of families. We applied LFSPRO to the family history "collected" at each time point. In this analysis, we calculated the 95% CIs for the differences in AUCs by first taking the differences in AUCs for each bootstrap sample and then calculating the means and 95%CIs.

Results

Clinical Illustration for counseling

Four scenarios of a hypothetical pedigree illustrate how LFSPRO estimates *TP53* mutation probability (Figure 1, Table 3). Figure 1A shows four family members diagnosed with early-onset cancers in each of four generations, suggesting a germline *TP53* mutation. The counselee is a potential mutation carrier by Chompret but not classic LFS criteria. Both criteria require the counselee to be a cancer patient. LFSPRO gives the counselee a carrier probability of 57% (Table 3), compared to 0.06% as observed by Myriad Genetics for a genetic testing population without knowledge of family history(35) (see Materials and Methods). Given the high mutation probability, when asymptomatic, the counselee has an elevated risk of developing cancer compared to the population: respectively 31% vs. 2% at age 40 years and 43% vs. 9% at age 70 years. In Figure 1B, the number and types of cancer are constant, but the status of the counselee's grandmother (sarcoma at age 25) is exchanged with that of the grandmother's sister (healthy until death from other causes at age 42). Early cancer in a grandmother within the counselee's direct lineage provides stronger evidence that the counselee may have inherited a germline mutation in *TP53*, and her carrier probability increases to 82%. By contrast,

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

the clinical prediction based on classic LFS or Chompret criteria, which use binary outcomes, is not sensitive to this change. In Figure 1C, more family members have cancer, further increasing the counselee's carrier probability to 91%, while the clinical criteria still do not recognize the family as LFS. In these three scenarios, the Mendelian transmission of a germline mutation in *TP53* likely occurred via the counselee's maternal line. Figure 1D demonstrates how LFSPRO captures a *de novo* mutation. Both of the counselee's maternal great-grandparents stayed healthy until death at an advanced age, while the counselee's mother had early-onset breast cancer, likely due to *de novo TP53* mutation. Because LFSPRO accounts for probabilities of *de novo* mutations, the counselee's carrier probability is 60%, rather than the 2% obtained from a strictly Mendelian model(24).

LFSPRO Predictive Performance

Three independent cohorts, comprising 867 families and 22,206 individuals, and 1,353 individuals tested for *TP53* mutations, were used to assess the performance of LFSPRO (Tables 1, 2) using six settings of the prevalence values for deleterious germline and *de novo TP53* mutations (Table S2). LFSPRO performed well in predicting the mutation carrier status under different scenarios of ascertainment. Incorporating *de novo* mutations into LFSPRO substantially improved its predictive power across all prevalence settings (Figures S3-S5). We selected one setting to demonstrate LFSPRO performance as follows. Clinical populations like our cohorts are likely to have enriched *TP53* mutations compared to that in the general population. The closest data we can reliably use to estimate a population undergoing genetic testing is that provided by Myriad Genetics(35), which showed the prevalence of germline mutations in *TP53* as 0.0006 (see Materials and Methods). We set the percentage of *de novo* mutations to be 20%(23), as a similar level of *de novo* mutations was observed in the MDACC dataset (unpublished data). We also use this setting as the default in the LFSPRO software if the user has not otherwise specified this value. Under this setting, the AUCs for the three

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

datasets, MDACC, ISKS and NCI, are 0.85 (95% CI: 0.75, 0.93), 0.67 (95% CI: 0.54, 0.79) and 0.82 (95% CI: 0.78, 0.86), respectively. The OE ratios are 1.35 (0.99, 1.80), 1.62 (1.03, 2.55) and 1.28 (1.17, 1.39), respectively (Table S2).

Figure 2a compares the predictive performance of classic LFS and Chompret criteria and LFSPRO (also Figures S6-S8). LFSPRO provides probabilities rather than binary outcomes, and generates an ROC curve with changing sensitivities and specificities as the probability cutoff changes. For the MDACC dataset, classic LFS criteria were highly specific at predicting *TP53* mutation carrier status, but had 12.5% sensitivity; whereas Chompret criteria had higher sensitivity (50%), with a slightly higher false positive rate (7%). LFSPRO achieved higher sensitivity and specificity than the Chompret and classic criteria (Figures S3, S6). For ISKS and NCI datasets, Chompret criteria and LFSPRO achieved similar sensitivity and specificity (Figures S4-S5, S7-S8). However, providing probabilities further allows LFSPRO to reach higher sensitivity while losing a little specificity by lowering the cutoff value (moving along the curve toward the upper right), or to reach higher specificity while losing a little sensitivity by increasing the cutoff value (moving toward the lower left).

The cutoffs for BRCA1/2 genetic counseling using BRCAPRO(31) typically lie between 0.05 and 0.2. In the MDACC data, the respective values of sensitivity and specificity for LFSPRO are 0.84 and 0.86 at a cutoff of 0.05; 0.78 and 0.89 at a cutoff of 0.1; and 0.63 and 0.94 at a cutoff of 0.2. We demonstrate LFSPRO's ability to classify individuals using a cutoff of 0.2, and how frequently it corrects the two criteria's classification of individuals as mutation carriers or non-carriers (Table 4). For the MDACC dataset, LFSPRO classifies more individuals correctly than both criteria. For the ISKS dataset, LFSPRO adds correct classifications of non-carriers (higher specificity) as compared with the Chompret criteria but not the classic LFS criteria. This is likely due to the low number of total mutation carriers observed in ISKS (Tables 1, 2). Classic LFS criteria's low sensitivity makes it difficult to detect *TP53* mutation carriers, but its high specificity

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

reduces false positives. For the NCI dataset, which contains the highest number of mutation carriers of the three datasets, LFSPRO added correct classifications of mutation carriers (higher sensitivity) but also reclassified some non-carriers incorrectly as carriers (reduced specificity), as compared with both criteria.

Evolution of family history in the pediatric cohort

To better mimic the application of LFSPRO when family data are collected in genetic counseling clinics, we determined the performance of LFSPRO over time following the initial diagnosis of the MDACC pediatric sarcoma dataset. This analysis accounted for the diagnosis of new cancer cases during follow-up, and estimated the predictive power at various time points. Figure 2B shows LFSPRO ROC curves for four time points while rolling forward. The AUCs for LFSPRO increased steadily over time, from 0.50 (no predictive power) at the date of ascertainment to 0.86 (strongly predictive) at the date of last contact. Table S4 shows that the differences in AUCs in 10-year intervals are greater than 0 and are statistically significant when the time of follow-up is more than 10 years. Importantly, for each time point, the Chompret criteria performed similarly to LFSPRO at each time point and then performed a little worse than LFSPRO at the last time point (Figure 2B). This result provides empirical evidence to the importance of collecting family history and following up with these families over the years, as is being practiced by genetic counselers at high-risk cancer clinics.

Discussion

We have developed a probabilistic prediction algorithm, LFSPRO, for *TP53* germline mutations and personalized cancer risk associated with LFS (Table S1 describes LFSPRO software). We demonstrated the strength of LFSPRO in risk prediction and clinical utility using datasets that represent varied ascertainment and biological mechanisms related to *TP53* and mixtures of families seen by clinical geneticists and genetic counselors. The associated stand-

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

alone R software is freely available at http://bioinformatics.mdanderson.org/main/LFSPRO, and in the widely used BayesMendel R package version 2.1-3, freely available at http://bcb.dfci.harvard.edu/bayesmendel/. The LFSPRO software contains functions for classic LFS and Chompret criteria so that users can obtain results from those criteria and LFSPRO for the same data.

Ideally, information on affected family members should improve the sensitivity of *TP53* mutation probability prediction by LFSPRO, while information about unaffected family members should lower high false positive rates. This was evident when we added *TP53*-negative families into our test cohorts. In many *TP53*-negative families, the proband and close relatives also had early-onset LFS-spectrum cancers. The high AUCs on multiple datasets support LFSPRO's ability to discriminate carriers and non-carriers. The OE ratios being close to 1 (with the 95% CIs covering 1 or very close to 1) further supports that the carrier probabilities estimated by LFSPRO are consistent with the truth (Table S2). We further performed subgroup analysis of OE, dividing the individuals into three risk groups, with estimated probabilities in low risk: 0-0.33; medium risk: 0.34-0.67; and high risk: 0.68-1 (Table S5). The OE ratios show good calibration for all risk subgroups in the MDACC dataset, and for the medium- and high-risk subgroups in the NCI dataset. We observed a slight underestimation of carrier probabilities in the low-risk subgroup in the NCI dataset. In the ISKS dataset, there is underestimation in the low-risk subgroup (statistically significant) and overestimation in the medium- and high-risk subgroups (not statistically significant).

The difference in ascertainment may account for the different in the performance levels that result when using the model on the various cohorts. LFSPRO performs best with the MDACC cohort (ascertainment through a specific tumor type and age, i.e., pediatric sarcoma) and the NCI cohort (ascertainment through a specific type of family history, i.e., families of LFS type), as *TP53* germline mutations are the dominant cause of cancer in both cohorts. Therefore, applying LFSPRO to families in high-risk clinics that are recruited in a similar way will help counsel the

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

family and provide further guidance on screening. In contrast, in the ISKS dataset, an adult sarcoma cohort, many sarcomas common in adults are rare in children, and some, such as well-differentiated or de-differentiated liposarcoma, myxoid liposarcoma and synovial sarcoma, are not typically associated with somatic or germline *TP53* mutations(40,41). The ISKS cohort is population-based so the family history of cancer is not taken into consideration when recruiting families. Many of these families did not display the classic LFS family history of cancer, rather subsequent sequencing identified a group of participants that carried germline *TP53* mutations(42). In this case, LFSPRO may still be used, but in conjunction with risk prediction models for other genes when possible (e.g., BRCAPRO(31) for BRCA1/2 mutations in early-onset breast cancer), in order to help identify portions of the phenotype that are caused by *TP53* mutations or other mutations, as people are increasingly identified as carrying *TP53* mutations by non-traditional methods.

The practice of testing for *TP53* mutations has changed over time and is not always 100% accurate. The MDACC dataset was tested more than 10 years ago. The NCI and ISKS datasets were tested more recently, and therefore some might have included deletions or duplications (del/dups). For the NCI dataset, some participants had testing performed prior to enrollment, which might not have included del/dups. We consider the deleterious del/dups in *TP53* to be rare and the difference in sensitivity across these datasets is likely to be small. All test results are expected to be 100% specific, with sensitivity between 0.95 and 0.99. We decided not to impose an arbitrary number on sensitivity of genetic testing in our analysis. Mathematically, if we did impose a sensitivity value less than 1, our ROC results would not be affected, but the OE ratios would be multiplied by 1/sensitivity.

Classic LFS and Chompret criteria start with the counselee being a cancer patient and are therefore not applicable for directly predicting a healthy individual's carrier status. They also provide a binary prediction, rather than a probability between 0 and 1. Hence, LFSPRO can provide complementary information on LFS risk assessment. Previous risk prediction models

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

like BRCAPRO also provide probabilities between 0 and 1, with cutoffs for BRCA testing chosen by genetic counselors. We project the cutoffs of LFSPRO will converge to a certain range as this prediction software is used together with clinical criteria in clinical settings. The LFSPRO software currently provides an individual's future risk of developing any invasive cancer. In the future, this will be expanded to predict the risk of developing a specific cancer type over the years. Both functions will be evaluated in large cohorts, preferably in conjunction with screening studies of LFS to further improve the practice of genetic counselors and the management of LFS families.

De novo TP53 germline mutations have high prevalence in LFS(23), occurring in 7%-20% of families. LFSPRO is the first Mendelian model that allows for *de novo* mutations in order to accommodate this relatively frequent event. We show this improved LFSPRO performance compared to that of a strict Mendelian model (Figures S3-S5). LFSPRO well discriminated mutation carriers from non-carriers (evaluated by AUCs) under six sets of assumed prevalence values for germline and *de novo TP53* mutations. Since these values likely depend on ascertainment, users can select these in the LFSPRO software to optimize performance under different clinical scenarios.

Family history changes over time. The long follow-up in the MDACC pediatric sarcoma cohort allows us to visualize the evolution of family history over time. This pediatric cohort did not look like LFS at the time of ascertainment by any clinical criteria, but slowly became more LFS-like over the years (Figure 2B). This observation supports the recognized importance among genetic counselors in ascertaining families over a period of years, especially when they are working with pediatric cancer syndromes.

As with BRCAPRO(25) and BOADICEA(43), algorithms that integrate clinical information to guide molecular testing in familial cancer may considerably impact clinical practice. As more clinical genetic testing with large gene panels becomes routine, LFSPRO will potentially aid in

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

the interpretation of genetic variants in *TP53* of unknown significance and prove useful in the management of Li-Fraumeni syndrome.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Acknowledgements

We thank Dr. Giovanni Parmigiani for incorporating LFSPRO in the BayesMendel R package.

To the memory of Fred Li.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

References

1. Li FP, Fraumeni JF, Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? Ann Intern Med 1969;71(4):747-52.
2. Li FP, Fraumeni JF, Jr. Rhabdomyosarcoma in children: epidemiologic study and identification of a familial cancer syndrome. J Natl Cancer Inst 1969;43(6):1365-73.
3. Li FP, Fraumeni JF, Jr., Mulvihill JJ, Blattner WA, Dreyfus MG, Tucker MA, et al. A cancer family syndrome in twenty-four kindreds. Cancer Res 1988;48(18):5358-62.
4. Nichols KE, Malkin D, Garber JE, Fraumeni JF, Jr., Li FP. Germ-line p53 mutations predispose to a wide spectrum of early-onset cancers. Cancer Epidemiol Biomarkers Prev 2001;10(2):83-7.
5. Olivier M, Goldgar DE, Sodha N, Ohgaki H, Kleihues P, Hainaut P, et al. Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype. Cancer Res 2003;63(20):6643-50.
6. Malkin D, Li FP, Strong LC, Fraumeni JF, Jr., Nelson CE, Kim DH, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990;250(4985):1233-8.
7. Eeles RA. Germline mutations in the TP53 gene. Cancer Surv 1995;25:101-24.
8. Hwang SJ, Lozano G, Amos CI, Strong LC. Germline p53 mutations in a cohort with childhood sarcoma: sex differences in cancer risk. Am J Hum Genet 2003;72(4):975-83.
9. Wu CC, Shete S, Amos CI, Strong LC. Joint effects of germ-line p53 mutation and sex on cancer risk in Li-Fraumeni syndrome. Cancer Res 2006;66(16):8287-92.
10. Wu CC, Strong LC, Shete S. Effects of measured susceptibility genes on cancer risk in family studies. Hum Genet 2010;127(3):349-57.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

11. Villani A, Tabori U, Schiffman J, Shlien A, Beyene J, Druker H, et al. Biochemical and imaging surveillance in germline TP53 mutation carriers with Li-Fraumeni syndrome: a prospective observational study. Lancet Oncol 2011;12(6):559-67.

12. LIFSCREEN : Evaluation of Whole Body MRI for Early Detection of Cancers in Subjects With P53 Mutation (Li-Fraumeni Syndrome) (LIFSCREEN). <https://clinicaltrials.gov/ct2/show/record/NCT01464086>.

13. Magnetic Resonance Imaging Screening in Li Fraumeni Syndrome (SIGNIFY). <https://clinicaltrials.gov/ct2/show/NCT01737255>.

14. A surveillance study utilizing whole body magnetic resonance imaging and other surveillance procedures in people with germ line cancer gene mutations to investigate the prevalence and incidence of investigable lesions. <https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12613000987763>.

15. Clinical and Genetic Studies of Li-Fraumeni Syndrome. <https://clinicaltrials.gov/ct2/show/NCT01443468>.

16. Cancer genetics and prevention, Li-Fraumeni whole body BRI study. . <http://www.dana-farber.org/Adult-Care/Treatment-and-Support/Treatment-Centers-and-Clinical-Services/Cancer-Genetics-and-Prevention-Program.aspx>.

17. National Comprehensive Cancer Network. Genetic/familial high risk assessment: breast and ovarian. Li Fraumeni syndrome management. 2016.

18. Anupindi SA, Bedoya MA, Lindell RB, Rambhatla SJ, Zelley K, Nichols KE, et al. Diagnostic Performance of Whole-Body MRI as a Tool for Cancer Screening in Children With Genetic Cancer-Predisposing Conditions. AJR Am J Roentgenol 2015;205(2):400-8.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

19. Birch JM, Hartley AL, Tricker KJ, Prosser J, Condie A, Kelsey AM, et al. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. Cancer Res 1994;54(5):1298-304.

20. Chompret A, Abel A, Stoppa-Lyonnet D, Brugieres L, Pages S, Feunteun J, et al. Sensitivity and predictive value of criteria for p53 germline mutation screening. J Med Genet 2001;38(1):43-7.

21. Tinat J, Bougeard G, Baert-Desurmont S, Vasseur S, Martin C, Bouvignies E, et al. 2009 version of the Chompret criteria for Li Fraumeni syndrome. J Clin Oncol 2009;27(26):e108-9; author reply e10.

22. Bougeard G, Renaux-Petel M, Flaman JM, Charbonnier C, Fermey P, Belotti M, et al. Revisiting Li-Fraumeni Syndrome From TP53 Mutation Carriers. J Clin Oncol 2015;33(21):2345-52.

23. Gonzalez KD, Buzin CH, Noltner KA, Gu D, Li W, Malkin D, et al. High frequency of de novo mutations in Li-Fraumeni syndrome. J Med Genet 2009;46(10):689-93.

24. Chen S, Wang W, Broman KW, Katki HA, Parmigiani G. BayesMendel: an R environment for Mendelian risk prediction. Stat Appl Genet Mol Biol 2004;3:Article21.

25. Euhus DM, Smith KC, Robinson L, Stucky A, Olopade OI, Cummings S, et al. Pretest prediction of BRCA1 or BRCA2 mutation by risk counselors and the computer model BRCAPRO. J Natl Cancer Inst 2002;94(11):844-51.

26. Chen S, Wang W, Lee S, Nafa K, Lee J, Romans K, et al. Prediction of germline mutations and cancer risk in the Lynch syndrome. JAMA 2006;296(12):1479-87.

27. Wang W, Chen S, Brune KA, Hruban RH, Parmigiani G, Klein AP. PancPRO: risk assessment for individuals with a family history of pancreatic cancer. J Clin Oncol 2007;25(11):1417-22.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

28. Wang W, Niendorf KB, Patel D, Blackford A, Marroni F, Sober AJ, et al. Estimating CDKN2A carrier probability and personalizing cancer risk assessments in hereditary melanoma using MelaPRO. Cancer Res 2010;70(2):552-9.

29. Gonzalez KD, Noltner KA, Buzin CH, Gu D, Wen-Fong CY, Nguyen VQ, et al. Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations. J Clin Oncol 2009;27(8):1250-6.

30. Elston RC, Stewart J. A General Model for the Genetic Analysis of Pedigree Data. Human Heredity 1971;21:20.

31. Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. Am J Hum Genet 1998;62(1):145-58.

32. Katki HA, Blackford A, Chen S, Parmigiani G. Multiple diseases in carrier probability estimation: accounting for surviving all cancers other than breast and ovary in BRCAPRO. Stat Med 2008;27(22):4532-48.

33. Life Table by National Cancer Institute Surveillance, Epidemiology, and End Results Program (2008-2010) <http://seer.cancer.gov>.

34. Lalloo F, Varley J, Ellis D, Moran A, O'Dair L, Pharoah P, et al. Prediction of pathogenic mutations in patients with early-onset breast cancer by family history. Lancet 2003;361(9363):1101-2.

35. Rich T LM, Kidd J, Saam J, Lancaster J. Characterization of Li-Fraumeni Syndrome using a 25-gene hereditary cancer panel. Poster session presented at: San Antonio Breast Cancer Symposium 2015 Dec 8-12; San Antonio,TX.

36. Ambry Genetics. Ambry Genetics shows expanded TP53 testing identifies more carriers, raising questions about clnical definition of Li-Fraumeni syndrome. http://www.ambrygen.com/press- releases/ambry-genetics-shows-expanded-tp53-testing-identifies-more-carriers-raising-questions.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

37. Strong LC, Williams WR. The genetic implications of long-term survival of childhood cancer. A conceptual framework. Am J Pediatr Hematol Oncol 1987;9(1):99-103.

38. Bondy ML, Lustbader ED, Strom SS, Strong LC. Segregation analysis of 159 soft tissue sarcoma kindreds: comparison of fixed and sequential sampling schemes. Genet Epidemiol 1992;9(5):291-304.

39. Lustbader ED, Williams WR, Bondy ML, Strom S, Strong LC. Segregation analysis of cancer in families of childhood soft-tissue-sarcoma patients. Am J Hum Genet 1992;51(2):344-56.

40. Mitchell G, Ballinger ML, Wong S, Hewitt C, James P, Young MA, et al. High frequency of germline TP53 mutations in a prospective adult-onset sarcoma cohort. PLoS One 2013;8(7):e69026.

41. Ito M, Barys L, O'Reilly T, Young S, Gorbatcheva B, Monahan J, et al. Comprehensive mapping of p53 pathway alterations reveals an apparent role for both SNP309 and MDM2 amplification in sarcomagenesis. Clin Cancer Res 2011;17(3):416-26.

42. Ballinger ML, Goode DL, Ray-Coquard I, James PA, Mitchell G, Niedermayr E, et al. Monogenic and polygenic determinants of sarcoma risk: an international genetic study. Lancet Oncol 2016;17(9):1261-71.

43. Antoniou AC, Pharoah PP, Smith P, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. Br J Cancer 2004;91(8):1580-90.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Tables

Table 1. Summary of demographic information for families from three study cohorts.

	MDACC Pediatric Sarcoma		ISKS Adult Sarcoma		NCI LFS		Total
Among All Families	No. of carriers>0	No. of carriers=0	No. of carriers>0	No. of carriers=0	No. of carriers>0	No. of carriers=0	
No. of families	11 (6.0%)	172 (94.0%)	19 (3.3%)	563 (96.7%)	78 (76.5%)	24 (23.5%)	867
No. of individuals	1,256 (49.2%)	1,297 (50.8%)	591 (3.5%)	16,386 (96.5%)	1,857 (69.4%)	819 (30.6%)	22,206
Average age at first diagnosis	46.4 (SD=23.9)	26.2 (SD=24.3)	46.3 (SD=21.8)	54.7 (SD=18.7)	39.9 (SD=22.1)	52.5 (SD=22.0)	-
No. of individuals with no. of cancers = 1	127 (10.1%)	265 (20.4%)	78 (13.2%)	1,788 (10.9%)	336 (18.1%)	159 (19.4%)	2,753
No. of individuals with No. of cancers = 2	19 (1.5%)	0 (0%)	7 (1.2%)	165 (1.0%)	91 (4.9%)	22 (2.7%)	304
No. of individuals with No. of cancers > 2	7 (0.6%)	0 (0%)	8 (1.4%)	48 (0.3%)	51 (2.7%)	11 (1.3%)	125

No.: number; SD: standard deviation

Table 2. Summary of individuals tested for *TP53* mutation status from three study cohorts.

	MDACC Pediatric Sarcoma		ISKS Adult Sarcoma		NCI LFS		Total
Among *TP53* Tested Individuals	Carriers	Non-carriers	Carriers	Non-carriers	Carriers	Non-carriers	Total
No. of individuals	32 (11.4%)	249 (88.6%)	25 (4.1%)	588 (95.9%)	210 (45.8%)	249 (54.2%)	1,353
Average age at first diagnosis	25.4 (SD=24.4)	13.0 (SD=12.7)	34.4 (SD=14.3)	46.0 (SD=17.7)	29.0 (SD=16.0)	42.1 (SD=15.3)	-

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Table 3. Clinical illustration of LFSPRO.

		Counselee's Absolute Risk of Developing Cancer [b]			
	Carrier Probability	By Age 40 years	By Age 70 Years	Classic Criteria	Chompret Criteria
From a baseline population[34]	0.06%	2%	9%	-	-
Shown in Figure 1A (scenario 1)	57%	31%	43%	N	Y
Exchange the status of counselee's grandmother and the grandmother's sister (scenario 2, Figure 1B)	82%	40%	52%	N	Y
Counselee's grandmother is also affected at age 42 years (scenario 3, Figure 1C)	91%	43%	56%	N	Y
Only the counselee and her mother and grandmother are affected (scenario 4, Figure 1D)[a]	60% 2%*	31% 3%*	43% 10%*	Y	Y

[a] All carrier probabilities are calculated allowing for de novo mutations, except for scenario 4, where we calculated the carrier probability (*) at de novo mutation rate of 0.

[b] Assuming the counselee has not developed cancer before age 35 years.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Table 4. Reclassification of *TP53* mutation carriers using LFSPRO.

	LFSPRO	Carriers		Non-carriers		Reclassification Rate[c]
		Pr≥0.2	Pr<0.2	Pr≥0.2	Pr<0.2	
MDACC	Classic					
	+	4	0[b]	0	1[a]	11.0%
	-	16[a]	12	14[b]	234	
	Chompret					
	+	11	3[b]	3	11[a]	12.1%
	-	9[a]	9	11[b]	224	
ISKS	Classic					
	+	2	0[b]	1	0[a]	2.6%
	-	4[a]	19	12[b]	575	
	Chompret					
	+	3	7[b]	2	32[a]	8.6%
	-	3[a]	12	11[b]	543	
NCI	Classic					
	+	23	0[b]	1	0[a]	49.7%
	-	146[a]	41	82[b]	166	
	Chompret					
	+	94	18[b]	6	8[a]	38.8%
	-	75[a]	23	77[b]	158	

[a] LFSPRO corrected classifications of carrier or non-carrier, as compared to clinical criteria.

[b] The clinical criteria corrected classifications of carrier or non-carrier, as compared to LFSPRO

[c] Reclassification rate is defined as the percentage of individuals that were classified differently by LFSPRO (categories a and b) as compared with the other criteria.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

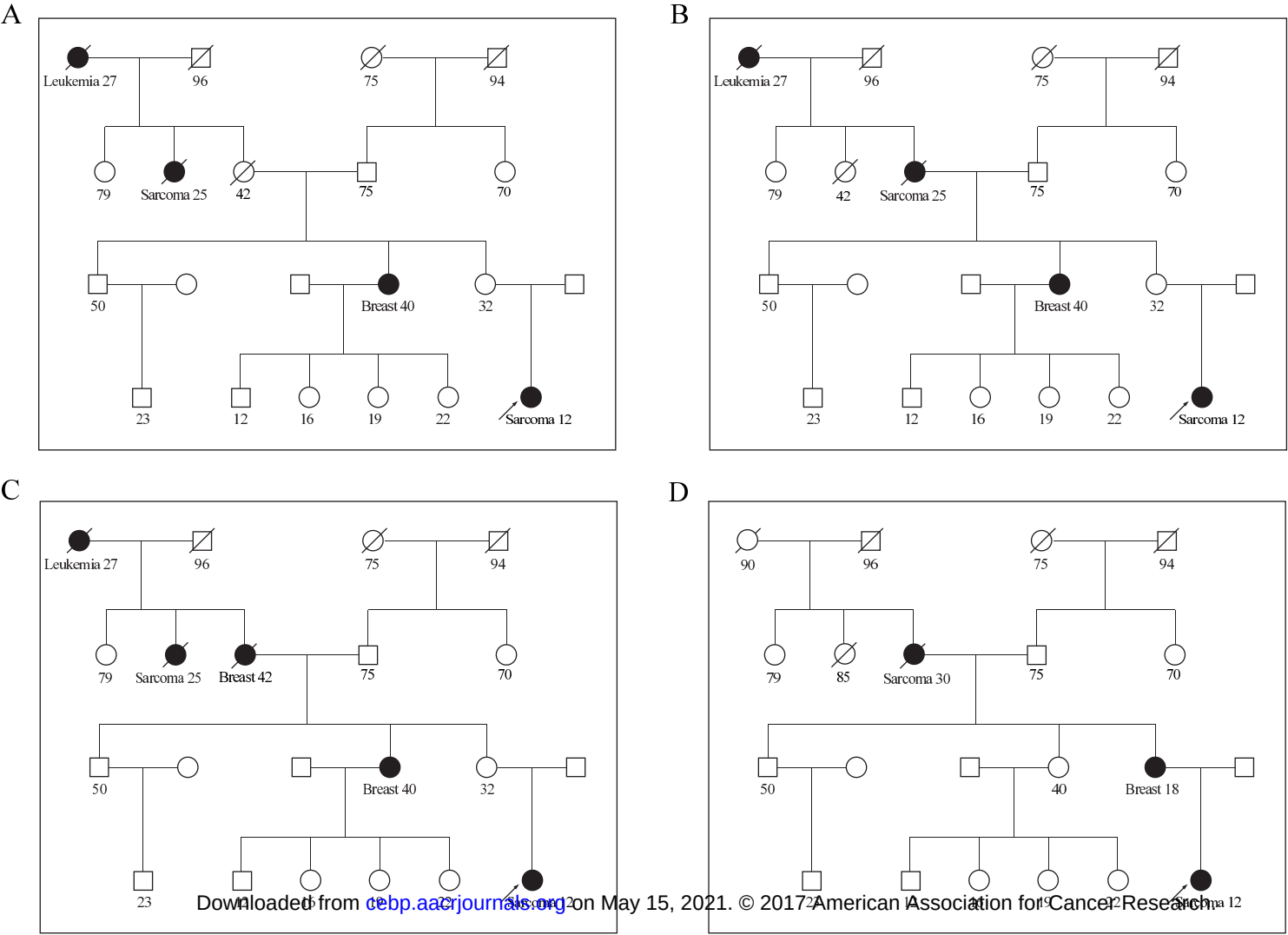
Figure Legends

Figure 1. A hypothetical family pedigree to illustrate the utility of LFSPRO for genetic counseling. An arrow indicates the counselee, for whom the *TP53* mutation probability is calculated by LFSPRO on the basis of the family history of cancer. Cancer type and age at diagnosis (years) in affected family members are indicated. (A-D) Four variations in clinical scenarios.

Figure 2. Validation results. (A) ROC curves of LFSPRO for three datasets. Also shown are true positive and false positive rates of classic LFS and Chompret criteria for the three datasets. (B) ROC curves of the MDACC dataset at different times of follow-up to evaluate the effect of ascertainment of family history in real clinical settings. The corresponding result from the Chrompret criteria is also shown. RF: roll forward. LCD: last contact date. In this analysis, the prevalence of germline mutation in *TP53* is 0.0006 and the percentage of *de novo* mutations among all *TP53* germline mutations is 20%.

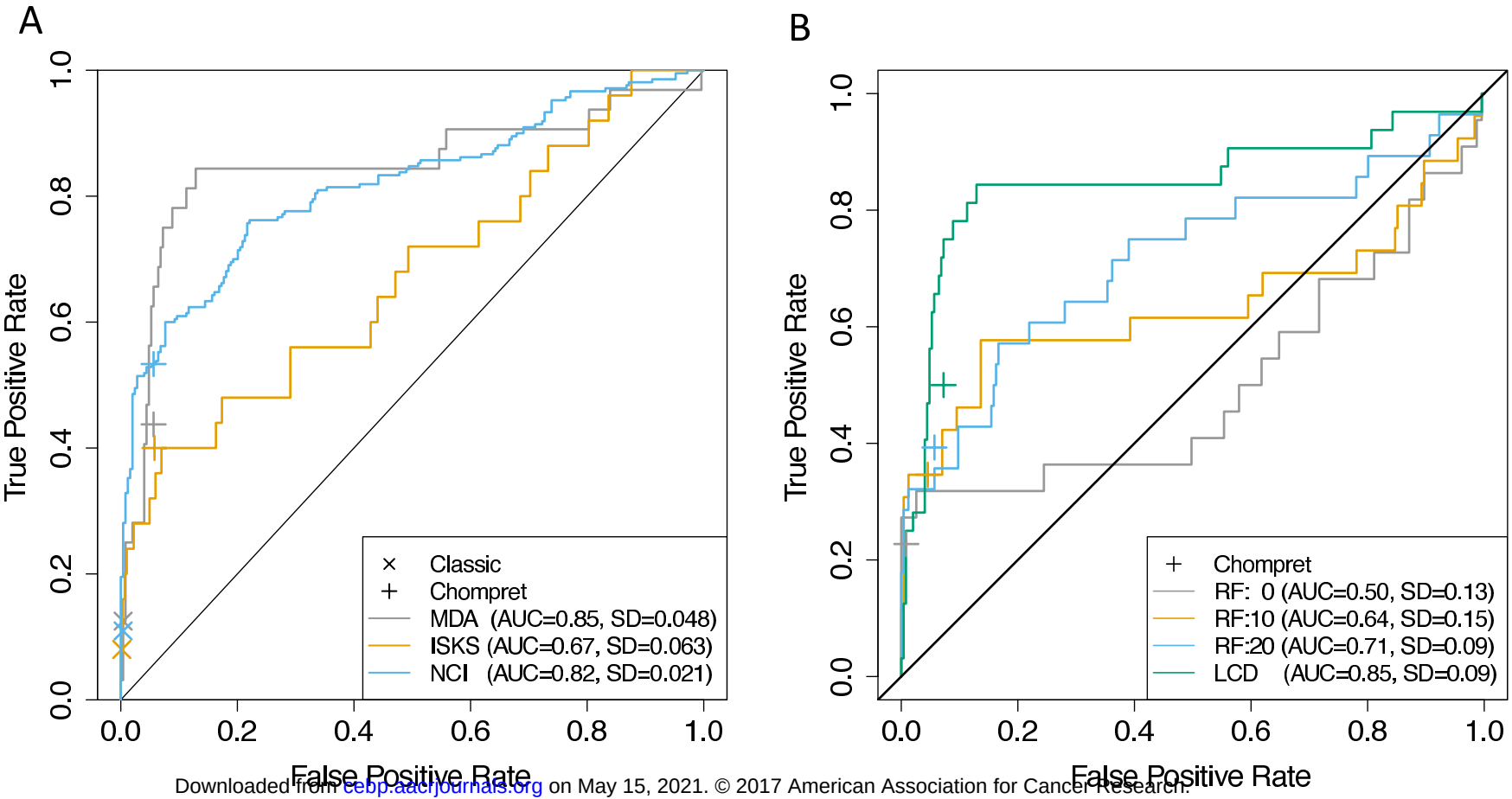
Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Figure 1



Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Figure 2



Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.

Author Manuscript Published OnlineFirst on January 30, 2017; DOI: 10.1158/1055-9965.EPI-16-0695
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Estimating TP53 Mutation Carrier Probability in Families with Li-Fraumeni Syndrome Using LFSPRO

Gang Peng, Jasmina Bojadzieva, Mandy L. Ballinger, et al.

Cancer Epidemiol Biomarkers Prev Published OnlineFirst January 30, 2017.

Updated version	Access the most recent version of this article at: doi:10.1158/1055-9965.EPI-16-0695
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2017/01/28/1055-9965.EPI-16-0695.DC1
Author Manuscript	Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

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, use this link http://cebp.aacrjournals.org/content/early/2017/01/28/1055-9965.EPI-16-0695. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.

Downloaded from cebp.aacrjournals.org on May 15, 2021. © 2017 American Association for Cancer Research.