Published in final edited form as:

Hum Genet. 2015 May ; 134(5): 497–507. doi:10.1007/s00439-014-1470-0.

Predicting survival in head and neck squamous cell carcinoma from *TP53* mutation

David L. Masica,
Department of Biomedical Engineering, Institute for Computational Medicine, The Johns Hopkins University, Baltimore, MD, USA

Shuli Li,
Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute, Boston, MA, USA

Christopher Douville,
Department of Biomedical Engineering, Institute for Computational Medicine, The Johns Hopkins University, Baltimore, MD, USA

Judith Manola,
Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute, Boston, MA, USA

Robert L. Ferris,
Department of Otolaryngology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Barbara Burtness,
Department of Medical Oncology, Fox Chase Cancer Center, Philadelphia, PA, USA

Arlene A. Forastiere,
Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Wayne M. Koch,
Department of Otolaryngology-Head and Neck Surgery, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

Christine H. Chung, and
Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Rachel Karchin
Department of Biomedical Engineering, Institute for Computational Medicine, The Johns Hopkins University, Baltimore, MD, USA. Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

David L. Masica: david.masica@gmail.com; Rachel Karchin: karchin@jhu.edu

Abstract

Correspondence to: Rachel Karchin, karchin@jhu.edu.

Electronic supplementary material The online version of this article (doi: 10.1007/s00439-014-1470-0) contains supplementary material, which is available to authorized users.

For *TP53*-mutated head and neck squamous cell carcinomas (HNSCCs), the codon and specific amino acid sequence change resulting from a patient's mutation can be prognostic. Thus, developing a framework to predict patient survival for specific mutations in *TP53* would be valuable. There are many bioinformatics and functional methods for predicting the phenotypic impact of genetic variation, but their overall clinical value remains unclear. Here, we assess the ability of 15 different methods to predict HNSCC patient survival from *TP53* mutation, using *TP53* mutation and clinical data from patients enrolled in E4393 by the Eastern Cooperative Oncology Group (ECOG), which investigated whether TP53 mutations in surgical margins were predictive of disease recurrence. These methods include: server-based computational tools SIFT, PolyPhen-2, and Align-GVGD; our in-house POSE and VEST algorithms; the rules devised in Poeta et al. with and without considerations for splice-site mutations; location of mutation in the DNA-bound TP53 protein structure; and a functional assay measuring *WAF1* transactivation in *TP53*-mutated yeast. We assessed method performance using overall survival (OS) and progression-free survival (PFS) from 420 HNSCC patients, of whom 224 had *TP53* mutations. Each mutation was categorized as "disruptive" or "non-disruptive". For each method, we compared the outcome between the disruptive group vs. the non-disruptive group. The rules devised by Poeta et al. with or without our splice-site modification were observed to be superior to others. While the differences in OS (disruptive vs. non-disruptive) appear to be marginally significant (Poeta rules + splice rules, $P = 0.089$; Poeta rules, $P = 0.053$), both algorithms identified the disruptive group as having significantly worse PFS outcome (Poeta rules + splice rules, $P = 0.011$; Poeta rules, $P = 0.027$). In general, prognostic performance was low among assessed methods. Further studies are required to develop and validate methods that can predict functional and clinical significance of *TP53* mutations in HNSCC patients.

Introduction

Cancer arises from the accumulation of genetic alterations selected to confer a growth advantage relative to healthy tissue (Vogelstein et al. 2013). Because these alterations impact cancer progression, there is much effort to isolate prognostic genetic markers (Ludwig and Weinstein 2005). *TP53* is among the most commonly altered genes in human cancers, and is found altered in high frequency across many diverse cancer types (Muller and Vousden 2013). This high prevalence positions *TP53* as a potentially powerful prognostic marker, if the differential impact of specific *TP53* mutations can be quantified. Indeed, a large body of work has investigated the use of *TP53*-based markers for cancer diagnosis, prognosis, and treatment (Olivier 2013; Robles and Harris 2010).

The prognostic value of *TP53* mutation appears to be highly context dependent. For instance, multiple independent groups have found that *TP53* mutation is associated with significantly reduced survival in chronic lymphocytic leukemia. (Dohner et al. 1995; Zenz et al. 2010; Rossi et al. 2009; Gonzalez et al. 2011) Conversely, multiple studies show no significant correlation between *TP53* status and survival in patients with glioblastoma. (Simmons et al. 2001; Newcomb et al. 1998; Weller et al. 2009) In a systemic review of 74 published studies of *TP53* status and survival in lung cancer patients, Steels et al. concluded that *TP53* alteration correlated with poorer survival in non-small-cell lung cancer, but data were insufficient to draw conclusions with regard to survival in small-cell lung cancer

(Steels et al. 2001). And, there is conflicting evidence about the role of *TP53* in ovarian carcinoma, including correlations with increased survival, decreased survival, and no observed correlation (Graeff 2009; HØGdall 2008; Darcy 2008; Ahmed et al. 2010; Bartel et al. 2008; Wang 2003); these different findings could arise from the methodologies used, small sample sizes, and mixed histological subtypes/grades (Wang 2003). Poeta et al. found a significant association between the presence of *TP53* mutation and decreased survival in head and neck squamous cell carcinoma (HNSCC) (Poeta et al. 2007). The authors further classified *TP53* mutations as "disruptive" or "non-disruptive", and found an even greater correlation between *TP53* mutations classified as disruptive and decreased survival; mutations classified as non-disruptive and survival were not significantly correlated. This implies that biomarker development can benefit from the classification of specific mutation, rather than treating all *TP53* mutations as causative. This is reasonable, because different *TP53* mutations are known to have a differential impact on carcinogenesis. (Bisio et al. 2014)

Many methods exist for evaluating the impact of specific genetic mutation on protein function. When assessed across large-scale multi-gene/multi-disease databases, popular computational methods typically achieve classification accuracies of ~60 % to ~80 % (Thusberg et al. 2011; Chan et al. 2007; Shihab et al. 2013). Unfortunately, classifier performance can vary significantly between specific genes and diseases, for a given method (Thusberg et al. 2011; Chan et al. 2007; Shihab et al. 2013; Hicks et al. 2011). This could result because methods are often developed and benchmarked for generalizability, and are agnostic to specific phenotype. This convention is important, because the probability that a mutation is "functional" or "deleterious" does not necessarily account its role in the context of a specific disease or cancer subtype. CHASM is a successful computational approach for identifying mutations that "drive" cancer (Carter et al. 2009; Integrated genomic analyses of ovarian carcinoma (2011); Parsons et al. 2011), and the authors of Align-GVGD provide highly curated sequence alignments for cancer-related genes (Tavtigian et al. 2008); however, even these methods do not necessarily discriminate distinct cancer-related phenotypes (e.g., survival, drug response, etc.). And, supervised learning approaches that classify the effect of mutation on specific phenotypes can be limited when training data are insufficient for the phenotype under consideration. In the era of genomic medicine and "big data", computational approaches for assessing the impact of mutation are indispensable, but the context in which they are applied can influence performance.

In this study, we classified *TP53* mutation as disruptive or non-disruptive using 15 different methods, and compared classification with survival in HNSCC patients harboring those mutations. The tested methods utilize diverse criteria for predicting the relative severity of mutations, including sequence, structure, and annotation, as well as an experimental assay of *TP53* function. These results contribute to the evolving role of *TP53* mutation as a potential prognosticator of HNSCC patient survival.

Materials and methods

We classified *TP53* mutation as "disruptive" or "non-disruptive" using 15 different classification approaches, and compared classification with overall and progression-free

survival in HNSCC patients harboring those mutations. All classifier predictions were submitted to the Eastern Cooperative Oncology Group (ECOG) to assess the prognostic power of each classifier. To control bias in a double-blind fashion, algorithm names were blinded before submission to ECOG and the authors that provided classification were blinded to all survival data, which are kept at the ECOG. Detailed information regarding the clinical trial E4393 was previously published (Poeta et al. 2007).

Classifiers

We classified *TP53* mutations as "disruptive" or "non-disruptive" using popular missense mutation prediction webservers, in-house algorithms, TP53 protein structure, an experimental assay, and the rules devised in Poeta et al. (2007) with and without special consideration for splice sites. Some of these methods only classify missense mutations, and could not provide classification for other mutation types (e.g., stop codons or frame shifts). Therefore, to help make final assessments as comparable as possible, classification for any mutation other than missense defaulted to the "Poeta rules" (see below) (Poeta et al. 2007). Of the 224 patients with *TP53* mutation, 148 had missense mutations, which included 75 unique missense mutations. A brief description of each classifier follows.

Sorting intolerant from tolerant (SIFT) (Kumar et al. 2009)—SIFT calculates the probability of each of the 20 amino acids occurring at each position in a multiple sequence alignment (MSA); probabilities are normalized based on the most frequently observed amino acid at that position in the alignment. Variants with a normalized probability ≤0.05 are considered disruptive, with all other variants considered non-disruptive. For this study, we used the SIFT webserver (http://sift.jcvi.org/) with default parameters and the default MSA.

PolyPhen-2 (Adzhubei 2010)—PolyPhen-2 combines sequence- and structure-based features to predict the impact of amino acid substitution using a naïve Bayes classifier. Sequence-based features include identity and conservation at the position in an MSA, and location of the variant at a hypermutable site or in a Pfam domain. Structure-based features include residue burial and crystallographic B-factor. For this study, we used the Poly-Phen-2 webserver (http://genetics.bwh.harvard.edu/pph2/) with default parameters and MSA. PolyPhen-2 classifies variants in one of three categories: benign, possibly damaging, and probably damaging. We submitted two independent sets of PolyPhen-2 predictions for ECOG assessment: (1) possibly damaging and probably damaging grouped together to define the disruptive class, and (2) probably damaging alone defines the disruptive class (all others considered non-disruptive).

Align grantham variation grantham deviation (Align-GVGD) (Tavtigian et al. 2008)—Align-GVGD considers biochemical variation at the relevant position in the MSA (Grantham Variation) and biochemical difference between reference and variant amino acids (Grantham Deviation). Variant predictions belong to one of seven classes (C0, C15, C25, C35, C45, C55, C65), where C0 is the least likely pathogenic and C65 is the most likely pathogenic. The server provides highly manually curated sequence alignments for more than a dozen cancer-related genes, including *TP53*; for this study, we selected the TP53 sequence

alignment with zebrafish as the most distant homolog (http://agvgd.iarc.fr/agvgd_input.php). We submitted two independent sets of Align-GVGD predictions for ECOG assessment: (1) C65, C55, and C45 classes define the disruptive group, and (2) C65, C55, C45, and C35 classes define the disruptive group (all others considered non-disruptive).

Variant effect scoring tool (VEST) (Carter et al. 2013; Douville et al. 2013)— VEST uses a supervised machine-learning algorithm, Random Forest, to identify missense mutations that are functional, i.e., they alter protein activity sufficiently to contribute to a clinically observable phenotype. Of all the tested methods, it uses the richest feature set, including properties of evolutionary conservation, biochemical properties of amino acid residues, predicted protein local structure, curated functional annotations, and protein sequence composition in the vicinity of a mutation. VEST reports a score for each mutation that ranges from 0 (least likely to be functional) to 1 (most likely to be functional). For this study, we used the CRAVAT web server (http://cravat.us) v.2.0 (VEST v.1.1.0, SNVBox v. 2.0.0) to obtain scores. We submitted two independent sets of VEST predictions for ECOG assessment: (1) VEST score ≥0.4 define the disruptive group, and (2) VEST score ≥0.8 define the disruptive group.

Phenotype-optimized sequence ensembles (POSE) (Masica et al. 2012)—POSE is a supervised learning approach that isolates sequences, from an input MSA, that facilitate an optimal classification of variants of known impact. Input data are a pre-computed MSA, a list of mutations, and the classification of each input mutation; the classification can be binary (e.g., disease causing or benign) or continuous valued (e.g., time-to-event data). The optimized MSA that results from training is then used to predict the impact of mutation using the POSE score function. POSE score is based on amino acid and physiochemical conservation derived from the MSA, and location of mutation in the target protein's 3D structure, if available. For this study, POSE was trained using continuous valued, mutation-specific *WAF1* transactivation (see below). From a database of 2,314 *TP53* mutations with associated *WAF1* transactivation, we selected 487 mutations to train POSE. Mutations used for training were associated with ≤20 % wild-type *WAF1* transactivation and observed at least 10 times in cancer-specific databases (putatively cancer-causing class), or were associated with ≥50 % wild-type activity and found fewer than three times in cancer-specific databases (putatively benign class). A DNA-bound structure of TP53 was used for training and prediction (PDB ID: 3EXJ).

Rules devised in Poeta et al. (2007)—Poeta et al. devised a simple rule-based method to classify a *TP53* mutation as disruptive or non-disruptive in the context HNSCC patient survival. Disruptive variants are those that introduce a stop codon anywhere in the gene, or any mutation in the L2 or L3 DNA-binding domains (codons 163-195 or 236-251) that replaces an amino acid from one charge/polarity category with an amino acid from different charge/polarity category (Poeta et al. 2007); all other variants are classified as non-disruptive.

Classification of TP53 splice sites—All variants were mapped to the *TP53* transcript ENST00000269305. The percent change in splice-site strength caused by the mutation was

calculated using Maximum Entropy Scan (Yeo and Burge 2004) and Human Splice Finder (Desmet 2009). Mutations causing significant changes in splice-site strength are likely to produce aberrant splicing and therefore not produce a functional protein. A change greater than 20 % for Maximum Entropy Scan (Bonnet et al. 2008) and 10 % for Human Splice Finder (Desmet 2009) will likely produce aberrant splicing.

The dominant isoform of *TP53* requires all the exons in which splice-site mutations were observed. The dominant isoform is the only transcript to fully express the full N-terminal transactivation domain, which is crucial for *TP53* function (Marcel et al. 2011). Even if other isoforms were produced as a result of the damaging splice mutations, the other isoforms will have differing functions from the dominant isoform.

Classification derived from the TP53 protein structure—We applied two simple structure-based rules, separately, to classify TP53 mutations using a DNA-bound TP53 crystal structure (PDB ID: 3EXJ). First we considered residue Euclidian distance to the bound DNA, which is a known correlate of carcinogenesis (Gentile et al. 1999; Lothe et al. 1995; Powell et al. 2000); two sets of predictions were submitted using the distance criteria: (1) any residue ≤20 Å or (2) any residue ≤10 Å to the bound DNA were defined as disruptive with all others defined as non-disruptive. Distances were calculated from closest residue–DNA atomic pair, for each residue. This calculation only considered the site of mutation, not the specific amino acid substitution.

Second, we investigated whether disruptive mutations cluster together in the TP53 protein structure more than non-disruptive mutations. This was accomplished by calculating the pairwise residue–residue distance for all sites that had a missense mutation, and counting the number of neighboring sites of mutation within 10 Å; distances were calculated from each residue's center of geometry (based on atomic coordinates). A relative clustering density was calculated by normalizing all sites relative to the site with the greatest number of counts (i.e., max clustering density = 1.0, with decreasing values corresponding to decreasing clustering density). This calculation only considered the site of mutation, not the specific amino acid substitution. All sites of mutation with a clustering density ≥0.5 were classified as disruptive, with all other sites classified as non-disruptive.

Classification derived from WAF1 transactivation—The WAF1 gene is thought to be transcriptionally activated by TP53 (El-Deiry 1993). WAF1 expression is induced in cells harboring wild-type *TP53*, but attenuated to varying degrees in *TP53*-mutated cells, including strong attenuation in cells harboring tumor-derived *TP53* mutations (El-Deiry 1993; Kato et al. 2003). Kato et al. (2003) quantified WAF1 transactivation in yeast for 2,314 individual TP53 amino acid substitutions, using a GFP-reporter assay. Here, we used the author-recommended cutoff of ≤20 % wild-type activity to define disruptive mutations, with all others defined as non-disruptive (Soussi et al. 2005). In addition, we submitted predictions with a more conservative cutoff of ≤10 % wild-type activity to define disruptive variants.

Assessment of model performance

Patient characteristics at baseline by mutation status were analyzed descriptively. Fisher's exact (Cox 1970) test or Mehta's exact test (Mehta et al. 1984) were used to evaluate the distribution of patient characteristics across mutation categories. Overall survival (OS) and progression-free survival (PFS) were estimated using the Kaplan–Meier method (Kaplan and Meier 1958). OS was defined as the time from study entry to death or to the last follow-up, and PFS was defined as the time from study entry to disease recurrence or death, whichever occurred first. Log-rank test (Peto and Peto 1972) and C-statistics (Uno et al. 2011) were used to evaluate each classifier univariately. Proportional hazard cox models (Cox 1972) were used to evaluate the significance of the following two classifiers: Poeta rules and the Poeta rules + Splice, after adjusting for important covariates.

Results

We classified 224 *TP53* mutations as "disruptive" (D) or "non-disruptive" (ND) using 15 different methods, and compared classification for overall survival (OS) and progression-free survival (PFS) in head and neck squamous cell carcinoma (HNSCC) patients. The supplementary Table S1 summarizes patient characteristics. Table 1 shows performance statistics for HNSCC OS, and classification distribution (D vs. ND), for each of the tested classifiers. Considering a significance threshold of $P \leq 0.05$, association between OS and disruptive vs. non-disruptive classification is not significant for any algorithm. Similarly, all hazard ratio 95 % confidence intervals include 1.0. And, C-statistics were low for all algorithms, showing classification accuracy slightly better than random. Seven of the classifiers, the rules devised in Poeta et al. with or without rules for splice-site mutations, both WAF1 classifiers, site-specific cluster density, PolyPhen-2 possibly damaging, and VEST ≥0.4 achieved a hazard ratio >1.0 (i.e., disruptive mutations are associated with reduced survival). Conversely, the other eight classifiers associate disruptive mutations with increased survival.

Figure 1 shows Kaplan–Meier survival curves for all seven classifiers that had hazard ratios (D vs. ND) >1.0, and are therefore consistent with the direction that disruptive mutations have reduced overall survival. For the Poeta rules with or without additional splice-site rules, the survival curves show reasonable separation between disruptive and non-disruptive classes (Fig. 1a, b), while for the remaining five classifiers, the curves are less well separated (Fig. 1c–g).

Table 2 shows performance statistics for HNSCC progression-free survival (PFS) for each of the tested classifiers. For PFS, only five classifiers estimated HR (D vs. ND) to be >1.0 (i.e., consistent with the direction that disruptive mutations to have decreased PFS). While the C-statistic remained low among all 15 algorithms, classification using Poeta rules with or without splice-site rules was significantly associated with patient PFS (HR (D vs. ND) = 1.43, 1.50; $P = 0.027$ and 0.011, respectively); none of the other classifiers reached significance. When evaluated in a multivariable setting adjusting for pathological T stage, N stage, and treatment, the mutation classifications using Poeta rules with or without splice-site rules remain significant for PFS (HR (D vs. ND) = 1.40, 1.46; $P = 0.040$, 0.020, respectively). Notably, Poeta rules with or without splice-site modification are the strongest

predictors of both OS and PFS, and these classifiers were more conservative with respect to classifying mutations as disruptive, compared with other classifiers (Table 1, final two columns). This result suggests that performance of the other 13 algorithms might suffer from over classifying mutations as disruptive.

Figure 2 shows Kaplan–Meier survival plots for all classifiers with hazard ratios (D vs. ND) >1 in the context of PFS. As with OS, classification via Poeta rules with or without splice-site rules results in survival curves with good separation between disruptive and non-disruptive classes (Fig. 2a, b). All five algorithms that had a hazard ratio >1 for PFS (Fig. 2) also had a hazard ratio >1 for OS (Fig. 1); however, classification using the WAF1 or cluster density classifiers was not statistically significant.

Next, we evaluated classifier performance on different subgroups of HNSCC patients, defined by: (1) newly diagnosed patients who underwent surgery only, (2) newly diagnosed patients who underwent surgery followed by postoperative therapy, (3) the surgery-only group combined with patients who underwent surgery followed by postoperative therapy (i.e., groups 1 and 2 combined), and (4) patients with recurrent disease who were treated with salvage surgery (Table 3). As with survival in the total study population, classification using the Poeta rules with or without splice-site rules resulted in the most significant correlation with survival in each of the patient subgroups, compared to the other 13 methods. Table 3 shows performance using these two classifiers for the patient subgroups, for both OS and PFS; performance statistics for the total study group are included for comparison, and are the same as in Tables 1 or 2.

For the surgery-only subgroup, including splice-site modification resulted in better correlation with both OS and PFS, compared with Poeta rules alone (Table 3). However, this correlation was only significant for PFS. Similarly, including splice-site modification resulted in significant correlation with PFS when the surgery only and surgery-plus postoperative therapy patients were grouped together; no significant correlation was observed for this grouping using the Poeta rules alone. And, classification was not significant for predicting survival in the surgery-plus postoperative-therapy subgroup, regardless of the classifier employed. Taken together, the observations from Table 3 indicate that the additional splice-site rules provide their greatest utility for predicting PFS, and treatment variation can influence the correlation.

Comparing hazard ratios and P values among the different subgroups, it is clear that the greatest correlation between classification and survival occurred in the recurrent subgroup (Table 3). Classification with either of the two algorithms was significantly correlated with PFS in the recurrent (salvage surgery) subgroup (HR = 2.12, 2.08; $P = 0.031, 0.038$, respectively). Classification using the Poeta rules also had significant correlation with OS in the salvage surgery subgroup (HR = 2.50 and P value = 0.015). Conversely, classification using the Poeta + splice-site rules was not significantly correlated with OS in the recurrent subgroup ($P = 0.081$).

Finally, we sought to the better understand the general discordance observed between survival and classifier predictions. The authors testing the classifiers were blinded to the

ECOG survival data, which are important for benchmarking classifiers and to prevent overfitting. To some extent, this limits our ability to determine the exact causes of the observed discordance. However, some potential origins of discordance can be inferred by comparing classification among the methods. Because Poeta-rule classification had higher correlation with survival than other tested classifiers, and have a distinct classification distribution (final column, Table 1), we compare all classifiers to the Poeta rules. Note that comparison with the splice-site rules is not relevant because no other classifier considered splice sites.

The IARC TP53 database is the "gold standard" database of *TP53* mutations (Petitjean et al. 2007). The IARC uses WAF1 ≤ 20 % as a criterion for defining *TP53* loss-of-function mutations. Table 4 provides a comparison of mutation classification between the WAF1 ≤20 % classifier and the other classifiers. Correlation was extremely high for all classifiers, with the possible exception of the Poeta rules (Table 4), suggesting that these algorithms are classifying *TP53* loss-of-function mutations as disruptive. Similarly, Table 5 shows mutations that were the greatest source of disagreement between Poeta rules and WAF1 ≤20 % classification. In all cases, Poeta rules classified the mutation as non-disruptive and WAF1 ≤20 % classified the mutation as disruptive. Table 5 also shows which other algorithms classified the mutations as disruptive. Importantly, each of these mutations is among the most frequently observed *TP53* somatic missense mutations in human cancers. Furthermore, they are frequent in the dataset used here, accounting for more than one-third of all missense mutations (53/148). Taken together, Tables 4 and 5 suggest that most classifiers tested here might suffer from identifying cancer-causing mutations, which does not necessarily correlate with reduced survival and treatment response.

Discussion

We classified *TP53* mutation as "disruptive" or "non-disruptive" using 15 different classification schemes to predict survival in 224 HNSCC patients harboring those mutations. Importantly, the tested approaches utilize diverse classifiers that include sequence, structure, annotation, or a mixture of these three. And, the *WAF1* classifiers rely entirely on an experimental functional assay and a threshold for defining a *TP53* mutation as disruptive or not. In general, classification of *TP53* mutation by the tested approaches was not highly correlated with HNSCC patient survival. Only the rules devised in Poeta et al. with or without additional splice-site classification resulted in statistically significant correlation. And, even these two classifiers had only marginally significant correlation, with low C-statistics and highly variable performance among the patient subgroups.

It is important to note that none of the classifiers tested here were developed specifically to predict the impact of mutation on survival. VEST, SIFT, Align-GVGD, and Poly-Phen-2 were developed to predict the impact of missense mutation, where "impact" describes broad categories such as "functional" or "deleterious". The POSE algorithm was developed, ostensibly, to predict any specific phenotype a user wishes to predict. However, there is no publically available dataset relating *TP53* mutation to HNSCC survival; here we trained POSE with mutation-specific *WAF1* transactivation data. The annotation- and structure-based classifiers tested utilize common rules for identifying mutations that might destabilize

protein structure or intermolecular interactions. Poeta rules utilize similar annotation/structure-based features, and also consider the physiochemical change imparted by the specific amino acid substitution. And, the *WAF1* transactivation assay was designed to predict loss-of-function mutation in *TP53*. Correlation between the *WAF1* loss-of-function assay and that of most other tested classifiers was highly significant (Table 4). Similarly, many of the tested methods classified the most canonical, cancer-causing *TP53* mutations as disruptive (Table 5); these mutations account for over one-third of the missense mutations in this study. Therefore, it appears to some extent, that the tested methods are working "as advertised" and identifying mutations that alter *TP53* function.

For some HNSCC patient subgroups, Poeta-rule classification was significantly associated with survival. Continued development of this classifier could result in improved prognostic markers for HNSCC patient survival. For instance, in this study we included new rules for classifying *TP53* splice-site mutations, and tested Poeta rules with or without the new splice-site classifier. Indeed, for some patient subgroups, the splice-site classifier added significant benefit. It may also be useful to add clinical variables for classification. As an example, Poeta-rule classification was most correlated with survival for the patient subgroup that underwent salvage resection. Based on the fact that these patients had recurrence, the set of mutations associated with these patients may have been enriched to predict poor survival. It is possible that Poeta-rule classification worked particularly well for these mutations by increasing the incidence within the group and improving the positive predictive value. Therefore, an improved classifier might include information about patient subgroup, such as newly diagnosed vs. recurrent, for classifying mutation as disruptive or non-disruptive.

Another possibility would be to train a classifier using HNSCC patient survival data. For proper benchmarking, and to prevent classifier overfitting, sufficiently large training and testing datasets would be required. This could include leave-some-out cross validation using a dataset as large as the ECOG dataset, which we used here solely for testing purposes. Another, perhaps more revealing strategy, would be to train and test using two independent clinical studies of HNSCC patient survival, where patient *TP53* sequence information was also available.

Potential confounding factors in our analysis include limitations in the sequencing assay, the absence of resected normal tissue adjacent to patient tumors, and absence of sequencing data from tumor subclones. In this cohort, insertion–deletion mutations greater than one base cannot be detected by the employed sequencing method. A patient with undetected *TP53* mutations could cause a false positive in our study, because the observed survival could be falsely attributed to that patient's detected mutation. In addition, *TP53* sequencing was restricted to the bulk of the tumor cells. This approach would not be able to detect mutations at the surgical margins or in subclones, which could be important because HNSCC heterogeneity can influence survival outcome. Similarly, our data do not provide any prognostic information regarding the presence of mutations in the adjacent normal mucosa due to the field cancerization (Bradley et al. 2007; Tabor et al. 2001).

It is also possible that there is a significant limitation to the use of single-gene markers. Cancer is complex disease progressing via the concerted impact of multiple genetic

alterations that can vary among individual tumors, even of similar histology. From this perspective, genetic context could affect the prognostic value of a single-gene marker. Furthermore, single-cell sequencing techniques are confirming that individual tumors are often heterogeneous mixtures of subclonal genotypes, where the competing populations can cause genotypic shifts over time (Greaves 2012; Marusyk et al. 2012; Keats et al. 2012). If the measured genotype is significantly different than that at later stages in progression, false inference could be drawn about the value of the biomarker under consideration. And of course, personal, environmental, and other exogenous factors can exacerbate an already poor genetic prognosis, which if not accounted for by the model can be confounding. To that end, improved prognostic biomarker selection will likely benefit from combining multiple genetic factors and clinical variables. Future prospective trials may be useful to determine whether particular mutations are predictive as well as prognostic, regarding use of cytotoxic chemotherapy or p53-directed immunotherapy (Ferris et al. 2014).

References

Adzhubei IA, et al. A method and server for predicting damaging missense mutations. Nat Meth. 2010; 7:248–249. http://www.nature.com/nmeth/journal/v7/n4/suppinfo/nmeth0410-248_S1.html.

Ahmed AA, et al. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. J Pathol. 2010; 221:49–56.10.1002/path.2696 [PubMed: 20229506]

Bartel F, et al. Both germ line and somatic genetics of the p53 pathway affect ovarian cancer incidence and survival. Clin Cancer Res. 2008; 14:89–96.10.1158/1078-0432.ccr-07-1192 [PubMed: 18172257]

Bisio A, Ciribilli Y, Fronza G, Inga A, Monti P. TP53 mutants in the tower of babel of cancer progression. Hum Mutant. 2014:n/a–n/a.10.1002/humu.22514

Bonnet C, et al. Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter mini-gene. J Med Genet. 2008; 45:438–446.10.1136/jmg.2007.056895 [PubMed: 18424508]

Bradley PJ, MacLennan K, Brakenhoff RH, Leemans CR. Status of primary tumour surgical margins in squamous head and neck cancer: prognostic implications. Current Opin Otolaryngol Head Neck Surg. 2007; 15:74–81.

Carter H, et al. Cancer-specific high-throughput annotation of somatic mutations: computational prediction of driver missense mutations. Cancer Res. 2009; 69:6660–6667.10.1158/0008-5472.can-09-1133 [PubMed: 19654296]

Carter H, Douville C, Stenson P, Cooper D, Karchin R. Identifying mendelian disease genes with the variant effect scoring tool. BMC Genom. 2013; 14:S3.

Chan PA, et al. Interpreting missense variants: comparing computational methods in human disease genes CDKN2A, MLH1, MSH2, MECP2, and tyrosinase (TYR). Hum Mutat. 2007; 28:683–693.10.1002/humu.20492 [PubMed: 17370310]

Cox, IDR. Analysis of binary data. Methuen and Company Ltd; London: 1970. p. 61-62.

Cox DR. Regression models and life tables. JR stat soc B. 1972; 34:187–220.

Darcy KM, et al. Associations between p53 overexpression and multiple measures of clinical outcome in high-risk, early stage or suboptimally-resected, advanced stage epithelial ovarian cancers: a Gynecologic Oncology Group study. Gynecologic Oncology. 2008; 111:487–495.10.1016/j.ygyno.2008.08.020 [PubMed: 18834621]

de Graeff P, et al. Modest effect of p53, EGFR and HER-2/neu on prognosis in epithelial ovarian cancer: a meta-analysis. Br J Cancer. 2009; 101:149–159. http://www.nature.com/bjc/journal/v101/n1/suppinfo/6605112s1.html. [PubMed: 19513073]

Desmet FOO, et al. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009; 37:e67.10.1093/nar/gkp215 [PubMed: 19339519]

Dohner H, et al. p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. Blood. 1995; 85:1580–1589. [PubMed: 7888675]

Douville C, et al. CRAVAT: cancer-related analysis of variants toolkit. Bioinformatics. 2013; 29:647–648.10.1093/bioinformatics/btt017 [PubMed: 23325621]

El-Deiry WS, et al. WAF1, a potential mediator of p53 tumor suppression. Cell. 1993; 75:817–825.10.1016/0092-8674(93)90500-P [PubMed: 8242752]

Ferris RL, et al. Phase I dendritic cell p53 peptide vaccine for head and neck cancer. Clin Cancer Res. 201410.1158/1078-0432.ccr-13-2617

Gentile M, Bergman Jungeström M, Olsen KE, Söderkvist P, Wingren S. p53 and survival in early onset breast cancer: analysis of gene mutations, loss of heterozygosity and protein accumulation. Eur J Cancer. 1999; 35:1202–1207.10.1016/S0959-8049(99)00121-5 [PubMed: 10615230]

Gonzalez D, et al. Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. J Clin Oncol. 2011; 29:2223–2229.10.1200/jco.2010.32.0838 [PubMed: 21483000]

Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012; 481:306–313. http://www.nature.com/nature/journal/v481/n7381/abs/nature10762.html-supplementary-information. [PubMed: 22258609]

Hicks S, Wheeler DA, Plon SE, Kimmel M. Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed. Hum Mutat. 2011; 32:661–668.10.1002/humu.21490 [PubMed: 21480434]

HØGdall EVS, et al. Distribution of p53 expression in tissue from 774 Danish ovarian tumour patients and its prognostic significance in ovarian carcinomas. APMIS. 2008; 116:400–409.10.1111/j.1600-0463.2008.00917.x [PubMed: 18452430]

Integrated genomic analyses of ovarian carcinoma. Nature. 2011; 474:609–615. http://www.nature.com/nature/journal/v474/n7353/abs/nature10166-f1.2.html-supplementary-information. [PubMed: 21720365]

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958; 53:457–481.

Kato S, et al. Understanding the function, Äìstructure and function, Äìmutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc Natl Acad Sci. 2003; 100:8424–8429.10.1073/pnas.1431692100 [PubMed: 12826609]

Keats JJ, et al. Clonal competition with alternating dominance in multiple myeloma. Blood. 2012; 120:1067–1076.10.1182/blood-2012-01-405985 [PubMed: 22498740]

Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009; 4:1073–1081. [PubMed: 19561590]

Lothe RA, et al. Deletion of 1p loci and microsatellite instability in colorectal polyps. Genes Chromosom Cancer. 1995; 14:182–188. [PubMed: 8589034]

Ludwig JA, Weinstein JN. Biomarkers in Cancer Staging, Prognosis and Treatment Selection. Nat Rev Cancer. 2005; 5:845–856. [PubMed: 16239904]

Marcel V, et al. Biological functions of p53 isoforms through evolution: lessons from animal and cellular models. Cell Death Differ. 2011; 18:1815–1824. [PubMed: 21941372]

Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer. 2012; 12:323–334. [PubMed: 22513401]

Masica DL, Sosnay PR, Cutting GR, Karchin R. Phenotype-optimized sequence ensembles substantially improve prediction of disease-causing mutation in cystic fibrosis. Hum Mutat. 2012; 33:1267–1274.10.1002/humu.22110 [PubMed: 22573477]

Mehta CR, Patel NR, Tsiatis AA. Exact significance testing to establish treatment equivalence with ordered categorical data. Biometrics. 1984:819–825. [PubMed: 6518249]

Muller PAJ, Vousden KH. p53 mutations in cancer. Nat Cell Biol. 2013; 15:2–8. [PubMed: 23263379]

Newcomb EW, et al. Survival of patients with glioblastoma multiforme is not influenced by altered expression of P16, P53, EGFR, MDM2 or Bcl-2 genes. Brain Pathol. 1998; 8:655–667.10.1111/j.1750-3639.1998.tb00191.x [PubMed: 9804374]

Olivier, M. *p53* in the Clinics. Hainaut, P.; Olivier, M.; Wiman, KG., editors. Vol. 8. Springer; New York: 2013. p. 127-146.

Parsons DW, et al. The genetic landscape of the childhood cancer medulloblastoma. Science. 2011; 331:435–439.10.1126/science.1198056 [PubMed: 21163964]

Petitjean A, et al. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. Hum Mutat. 2007; 28:622–629.10.1002/humu.20495 [PubMed: 17311302]

Peto R, Peto J. Asymptotically efficient rank invariant test procedures. JR Stat Soc. 1972; 135:185–206.

Poeta ML, et al. TP53 Mutations and Survival in Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2007; 357:2552–2561.10.1056/NEJMoa073770 [PubMed: 18094376]

Powell B, Soong R, Iacopetta B, Seshadri R, Smith DR. Prognostic significance of mutations to different structural and functional regions of the p53 gene in breast cancer. Clin Cancer Res. 2000; 6:443–451. [PubMed: 10690522]

Robles AI, Harris CC. Clinical outcomes and correlates of TP53 mutations and cancer. Cold Spring Harb Perspect Biol. 2010; 2:a001016.10.1101/cshperspect.a001016 [PubMed: 20300207]

Rossi D, et al. The Prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. Clin Cancer Res. 2009; 15:995–1004.10.1158/1078-0432.ccr-08-1630 [PubMed: 19188171]

Shihab HA, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden markov models. Hum Mutat. 2013; 34:57–65.10.1002/humu.22225 [PubMed: 23033316]

Simmons ML, et al. Analysis of complex relationships between age, p53, epidermal growth factor receptor, and survival in glioblastoma patients. Cancer Res. 2001; 61:1122–1128. [PubMed: 11221842]

Soussi T, Kato S, Levy PP, Ishioka C. Reassessment of the TP53 mutation database in human disease by data mining with a library of TP53 missense mutations. Hum Mutat. 2005; 25:6–17.10.1002/humu.20114 [PubMed: 15580553]

Steels E, et al. Role of p53 as a prognostic factor for survival in lung cancer: a systematic review of the literature with a meta-analysis. Eur Respir J. 2001; 18:705–719. [PubMed: 11716177]

Tabor MP, et al. Persistence of genetically altered fields in head and neck cancer patients: biological and clinical implications. Clin Cancer Res. 2001; 7:1523–1532. [PubMed: 11410486]

Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB. In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat. 2008; 29:1327–1336.10.1002/humu.20892 [PubMed: 18951440]

Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. Hum Mutat. 2011; 32:358–368.10.1002/humu.21445 [PubMed: 21412949]

Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med. 2011; 30:1105–1117.10.1002/sim.4154 [PubMed: 21484848]

Vogelstein B, et al. Cancer Genome Landscapes. Science. 2013; 339:1546–1558.10.1126/science.1235122 [PubMed: 23539594]

Weller M, et al. Molecular predictors of progression-free and overall survival in patients with newly diagnosed glioblastoma: a prospective translational study of the German Glioma network. J Clin Oncol. 2009; 27:5743–5750.10.1200/jco.2009.23.0805 [PubMed: 19805672]

Wong KK, et al. Poor survival with wild-type TP53 ovarian cancer? Gynecol Oncol. 2013; 130:565–569.10.1016/j.ygyno.2013.06.016 [PubMed: 23800698]

Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol. 2004; 11:377–394. [PubMed: 15285897]

Zenz T, et al. TP53 Mutation and Survival in Chronic Lymphocytic Leukemia. J Clin Oncol. 2010; 28:4473–4479.10.1200/jco.2009.27.8762 [PubMed: 20697090]

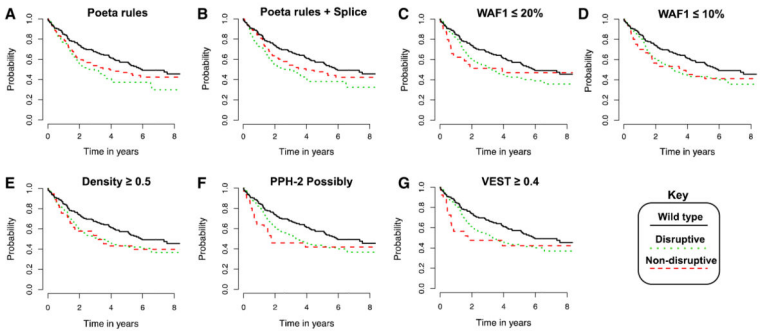

Fig. 1.

Kaplan-Meier survival curves for top-predicting classifiers of HNSCC overall survival. *Survival curves* for all classifiers that predicted a causative effect for disruptive mutations (i.e., hazard ratio >1.0) from Table 1. Overall survival among HNSCC patients with wild-type *TP53* shown for comparison

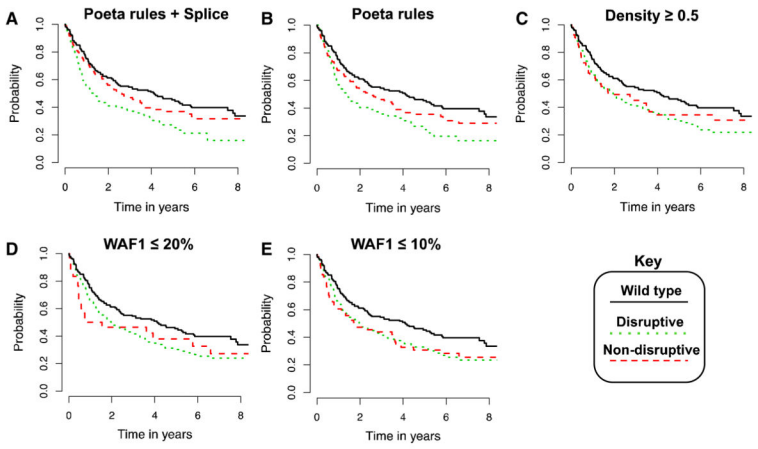

Fig. 2.

Kaplan-Meier survival curves for top-predicting classifiers of HNSCC progression-free survival. *Survival curves* for all classifiers that predicted a causative effect for disruptive mutations (i.e., hazard ratio >1.0) from Table 2. Progression-free survival among HNSCC patients with wild-type TP53 shown for comparison

Table 1

Classifying HNSCC overall survival from *TP53* mutation

Classifier	HR (D vs. ND)	HR, 95% CI	*P* value	*C*	*C*, 95% CI	Predictions	
						D, ND	D/(D + ND)
Poeta rules	1.40	1.0–1.98	0.05	0.55	0.51–0.60	85, 139	0.38
Poeta rules + Splice	1.34	0.96–1.89	0.09	0.55	0.51–0.60	105, 119	0.47
WAF1 ≤20 %	1.28	0.77–2.14	0.34	0.54	0.51–0.58	189, 35	0.84
WAF1 ≤10 %	1.12	0.78–1.62	0.54	0.54	0.51–0.58	151, 73	0.67
Density ≥0.5	1.09	0.75–1.60	0.65	0.54	0.51–0.58	160, 64	0.71
PPH-2 possibly	1.07	0.66–1.73	0.79	0.54	0.51–0.58	189, 35	0.84
VEST ≥0.4	1.07	0.62–1.82	0.82	0.54	0.50–0.58	195, 39	0.87
Distance ≤20 Å	0.99	0.66–1.47	0.95	0.54	0.50–0.58	169, 55	0.75
PPH-2 Probably	0.98	0.66–1.43	0.90	0.54	0.51–0.58	164, 60	0.73
Distance ≤10 Å	0.89	0.63–1.25	0.49	0.55	0.51–0.59	121, 103	0.54
VEST ≥0.8	0.86	0.57–1.29	0.47	0.55	0.50–0.59	177, 47	0.79
AGVGD C35	0.84	0.59–1.20	0.34	0.55	0.51–0.59	145, 79	0.65
SIFT	0.84	0.53–1.33	0.45	0.55	0.50–0.59	189, 35	0.84
AGVGD C45	0.79	0.56–1.11	0.18	0.55	0.51–0.59	130, 94	0.58
POSE	0.75	0.50–1.12	0.16	0.55	0.51–0.59	174, 50	0.78

Performance achieved by 15 different classifiers for predicting overall survival in HNSCC patients from *TP53* mutation. Classifiers include: the rules devised by Poeta et al. with or without splice-site considerations (Poeta rules + Splice and Poeta rules, respectively); site-specific clustering density ≥0.05 (Density ≥0.5); WAF1 transactivation as a percent of wild type (WAF1 ≤20 % and WAF1 ≤10 %); site of mutation—DNA distance at two discrete cutoffs (Distance ≤20 Å and Distance ≤10 Å); PolyPhen-2, with two different dichotomizations of the tripartite output (PPH-2 Possibly and PPH-2 Probably); the VEST algorithm with two different thresholds (VEST ≥0.8 and VEST ≥0.4); Align-GVGD, with two different dichotomizations of the septpartite output (AGVGD C45 and AGVGD C35); and the SIFT and POSE algorithms each with a single implementation

C is the concordance statistic, *CI* is the confidence interval, *HR* is the hazard ratio calculated for disruptive (D) vs. non-disruptive (ND) classifications, and Predictions shows the relative classification (D vs. ND) for the 230 TP53 variants, for each classifier

See “Materials and methods” for a complete description of each classifier and performance statistic

Table 2

Classifying HNSCC progression-free survival from *TP53* mutation

Classifier	HR (D vs. ND)	HR, 95 % CI	*P* value	*C*	*C*, 95 % CI
Poeta rules + Splice	1.50	1.10–2.04	0.01	0.57	0.53–0.60
Poeta rules	1.43	1.04–1.95	0.03	0.56	0.53–0.60
Density ≥0.5	1.09	0.77–1.54	0.63	0.55	0.52–0.59
WAF1 ≤20 %	1.07	0.69–1.65	0.76	0.55	0.52–0.58
WAF1 ≤10 %	1.06	0.76–1.47	0.74	0.55	0.52–0.58
Distance ≤20 Å	0.95	0.67–1.36	0.79	0.55	0.52–0.58
PPH-2 Probably	0.94	0.67–1.34	0.75	0.55	0.52–0.59
PPH-2 Possibly	0.94	0.62–1.44	0.79	0.55	0.52–0.58
Distance ≤10Å	0.90	0.66–1.22	0.48	0.55	0.52–0.59
VEST ≥0.4	0.88	0.55–1.39	0.57	0.55	0.52–0.59
VEST ≥0.8	0.86	0.60–1.25	0.44	0.55	0.52–0.59
AGVGD C35	0.83	0.60–1.14	0.24	0.56	0.52–0.59
AGVGD C45	0.80	0.58–1.09	0.15	0.56	0.52–0.59
SIFT	0.78	0.51–1.18	0.23	0.56	0.52–0.59
POSE	0.71	0.50–1.01	0.06	0.56	0.53–0.59

Performance achieved by 15 different classifiers for predicting progression-free survival in HNSCC patients from *TP53* mutation. See the Table 1 caption for a description of classifiers and performance statistics

Table 3

Classifying HNSCC progression-free survival (PFS) and overall survival (OS) for different HNSCC patient subgroups using the rules devised in Poeta et al. with and without splice-site consideration

Surgical group	Survival	Classifier	HR (D vs. ND)	HR, 95 % CI	*P* value
All patients with mutated *TP53* (*n* = 224)	PFS	Poeta rules + Splice	1.50	1.10–2.04	0.01
		Poeta rules	1.43	1.04–1.95	0.03
	OS	Poeta rules + Splice	1.34	0.96–1.89	0.09
		Poeta rules	1.40	1.00–1.98	0.05
Recurrent only (*n* = 39)	PFS	Poeta rules + Splice	2.12	1.07–4.22	0.03
		Poeta rules	2.08	1.04–4.16	0.04
	OS	Poeta rules + Splice	1.93	0.92–4.03	0.08
		Poeta rules	2.50	1.19–5.23	0.02
Surgery only (*n* = 69)	PFS	Poeta rules + Splice	1.98	1.10–3.57	0.02
		Poeta rules	1.72	0.93–3.17	0.08
	OS	Poeta rules + Splice	1.30	0.68–2.48	0.42
		Poeta rules	1.03	0.51–2.08	0.94
Surgery + postoperative therapy (*n* = 114)	PFS	Poeta rules + Splice	1.21	0.78–1.87	0.40
		Poeta rules	1.18	0.76–1.82	0.47
	OS	Poeta rules + Splice	1.28	0.79–2.09	0.31
		Poeta rules	1.42	0.88–2.30	0.16
Surgery only combined with surgery + postoperative therapy (*n* = 183)	PFS	Poeta rules + Splice	1.45	1.02–2.06	0.04
		Poeta rules	1.33	0.93–1.90	0.11
	OS	Poeta rules + Splice	1.28	0.87–1.89	0.21
		Poeta rules	1.26	0.85–1.86	0.25

HR is the hazard ratio calculated for disruptive (D) vs. non-disruptive (ND) classifications, *CI* is the confidence interval, and *n* is the number of patients in the corresponding subgroup

Results for All patients with mutated *TP53* are shown for comparison, and are the same as in Tables 1 or 2

Table 4

Correlation of predictions made by each classifier and the WAF1 ≤20 % classifier

Classifier	FDR-adjusted *P* value
Poeta rules	2.03×10^{-3}
Density ≥0.5	1.31×10^{-10}
Distance ≤20 Å	3.58×10^{-11}
Distance ≤10 Å	1.99×10^{-8}
PPH-2 possibly	3.58×10^{-11}
PPH-2 probably	3.58×10^{-11}
VEST ≥0.4	4.37×10^{-11}
VEST ≥0.8	3.77×10^{-11}
AGVGD C45	1.20×10^{-8}
AGVGD C35	5.22×10^{-10}
SIFT	3.58×10^{-11}
POSE	1.48×10^{-10}

The Poeta + splice and WAF1 ≤ 20 % predictions were not significantly co-occurring. The WAF1 ≤ 10 % classifier was not included in the calculation to prevent an artificially conservative FDR adjustment

Two-tailed Fisher's exact *P* values were calculated by populating a two-by-two contingency table with predictions made by each classifier and those made by the WAF1 ≤20 % classifier (disruptive = 1; non-disruptive = 0); *P* values were adjusted using the Benjamini and Hochberg false-discovery rate (FDR)

Table 5

Highest-frequency mutations from this cohort that were predicted disruptive using the WAF1 ≤20 % classifier and non-disruptive using Poeta rules

Missense mutation	Classified “disruptive” by:
R175H (8)	VEST, SIFT, PPH-2, POSE, density, distance
H193R (3)	VEST, SIFT, PPH-2, POSE, density, distance
Y205C (4)	VEST, SIFT, PPH-2, AGVGD, POSE
Y220C (8)	VEST, SIFT, PPH-2, AGVGD, POSE
G245S (3)	VEST, SIFT, PPH-2, AGVGD, POSE, density, distance
V272M (4)	VEST, SIFT, PPH-2, POSE, density, distance
R273C (8)	VEST, SIFT, PPH-2, AGVGD, POSE, density, distance
R273H (8)	VEST, SIFT, PPH-2, POSE, density, distance
R280K (3)	VEST, SIFT, PPH-2, POSE, density, distance
R282W (4)	VEST, SIFT, PPH-2, AGVGD, POSE, density, distance

Parenthetical values indicate the number of patients from this cohort with the mutation

For each mutation, algorithms that classified that mutation as disruptive are listed in the second column