diagnostics

MDPI

Article

Association of Fraction of Genome Altered with Overall Survival in Sarcoma: Histologic Subtype-Stratified Discovery and External Replication Study

Mehmet Mutlu Çatlı [1,*] and Mehmet Acun [2]

[1] Department of Medical Oncology, Antalya Training and Research Hospital, 07100 Antalya, Türkiye
[2] Department of Medical Oncology, Isparta City Hospital, 32200 Isparta, Türkiye
* Correspondence: catlimehmetmutlu@gmail.com; Tel.: +90-5542689623

Abstract

Background: Whole-genome doubling (WGD) is reported in clinical sequencing, but its prognostic value in sarcoma has not been assessed after accounting for histologic subtype. We evaluated whether fraction of genome altered (FGA) is associated with overall survival (OS) and whether WGD adds prognostic information beyond FGA. **Methods:** We analyzed 1494 MSK-IMPACT patients (593 deaths) across 17 subtypes using subtype-stratified Cox models adjusted for age and sample type; OS was measured from sequencing. Incremental value was assessed with likelihood-ratio tests and optimism-corrected C-indices. TCGA sarcomas (n = 241; 233 with ABSOLUTE-based WGD) were used for external replication. **Results:** FGA was associated with worse OS (hazard ratio [HR] per 0.10 increase, 1.11; 95% CI, 1.06 to 1.16; per unit increase, 2.84; 95% CI, 1.81 to 4.46), with a graded relationship across quartiles (HR per step, 1.21). WGD was associated with OS unadjusted (n = 1774: HR, 1.21; 95% CI, 1.02 to 1.43; n = 1494: HR, 1.25; 95% CI, 1.05 to 1.50) but not after subtype stratification (n = 1494: HR, 1.12; 95% CI, 0.93 to 1.36). Adding FGA to WGD improved model fit ($p < 0.001$); adding WGD to FGA did not ($p = 0.81$). Optimism-corrected C-indices were 0.618 (clinical model), 0.618 (plus WGD), and 0.633 (plus FGA). In TCGA, FGA was associated with OS (HR per unit, 5.15; 95% CI, 2.05 to 12.93) and added information beyond WGD ($p = 0.004$), whereas WGD added none beyond FGA ($p = 0.44$), although WGD alone remained associated with OS in TCGA (HR, 1.69). FGA effect size varied across subtypes ($I^2 = 57\%$) but persisted without gastrointestinal stromal tumors (HR, 2.06; 95% CI, 1.26 to 3.36). **Conclusions:** Higher FGA was associated with shorter OS in two sarcoma cohorts, and WGD added no detectable incremental information beyond FGA in either. Discrimination gains were modest; grade, stage, and treatment were not modeled in the main analyses; and OS was measured from sequencing. Prospective evaluation is needed before clinical use.

Keywords: genomic instability; copy-number alteration; survival analysis; prognostic biomarker; next-generation sequencing; histologic heterogeneity; precision oncology; external replication

Academic Editor: Marko Buta

Received: 26 September 2026
Revised: 4 November 2026
Accepted: 19 November 2026
Published: 25 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Genomic instability is regarded as an enabling characteristic of cancer, because it generates the genetic diversity on which tumor evolution depends [1]. One extreme form of instability is whole-genome doubling (WGD), in which a tumor cell duplicates its entire chromosome complement. Experimental work has suggested that WGD allows cells to tolerate chromosome missegregation and thereby favors further chromosomal instability [2]. In a clinical sequencing cohort of 9692 patients with advanced cancer, WGD was present in nearly 30% of tumors and was associated with worse outcomes independently of established clinical prognostic factors [3]. Aneuploidy more broadly has been linked to TP53 mutation and to the expression of proliferation-related genes across thousands of tumor genomes [4].

Sarcomas are a heterogeneous group of malignancies of soft tissue and bone that comprise many histologic subtypes [5]. Large-scale genomic studies have described their molecular features in detail [6,7]. However, most sarcoma subtypes are individually uncommon, and prognostic biomarkers are therefore often evaluated in mixed-histology cohorts. In such settings, a marker that differs in frequency between subtypes may appear prognostic largely because it reflects histologic composition.

This concern is relevant to WGD. The MSK-IMPACT assay reports copy-numpy-derived measures, including WGD status and the fraction of the genome altered (FGA), as part of routine clinical sequencing [8]. In a study of 2138 sarcomas spanning 45 pathologic entities, WGD rates varied across entities, and WGD was associated with shorter survival in metastatic, but not primary, tumors [9]. A recent secondary analysis of the same cohort combined WGD, FGA, and four other genomic features into a single composite instability score, framed explicitly as an alternative to histology-based risk stratification [10]; because the six features were combined into one score, that analysis did not address whether WGD and FGA provide independent or overlapping prognostic information. To our knowledge, no study has directly tested whether WGD provides prognostic information beyond FGA, or FGA beyond WGD, within a model that accounts for histologic subtype, although subtype-specific genomic models have been proposed for individual sarcoma subtypes, for example, malignant peripheral nerve sheath tumor [11]. WGD and FGA describe related but distinct features of the copy-number profile. FGA quantifies the extent of the genome affected by copy-number gain or loss, whereas WGD is an inferred genome-doubling event based on allele-specific copy-number patterns; WGD is therefore not a dichotomized form of FGA. Because FGA is analyzed here on its continuous scale, it avoids the loss of information that accompanies dichotomization of continuous variables [12], but any difference between the two markers in prognostic performance cannot be attributed to continuous versus binary measurement alone. Given that the two markers are related, it remains unclear whether FGA is associated with survival after accounting for histologic subtype, and how much prognostic information each marker carries conditional on the other.

We addressed these questions in a two-stage analysis. In the MSK-IMPACT sarcoma cohort, we used subtype-stratified Cox models to evaluate the association of FGA and WGD with overall survival and to test the incremental value of each marker over the other. We then examined whether the findings were replicated in an independent sarcoma cohort from The Cancer Genome Atlas [6]. Because this cohort differs in variables and in the survival time origin, we regard it as an external replication of the biomarker association and not as a validation of a fixed prediction model.

https://doi.org/10.3390/diagnostics16233974

2. Materials and Methods

2.1. Study Design and Data Sources

This was a retrospective analysis of two publicly available, de-identified sarcoma cohorts, reported in accordance with the REMARK recommendations for tumor marker prognostic studies [13]. The MSK-IMPACT sarcoma cohort [9] served as the discovery cohort, and the sarcoma cohort of The Cancer Genome Atlas (TCGA-SARC, PanCancer Atlas) served as the external replication cohort. Clinical and genomic data for both cohorts were obtained from cBioPortal [14,15] (study identifiers sarcoma_mskcc_2022 and sarc_tcga_pan_can_atlas_2018; downloaded in May and July 2026, respectively), and the genome-doubling calls for TCGA-SARC were obtained from the PanCancer Atlas ABSOLUTE call set in September 2026. Because only public, de-identified data were used, no patient contact or new data collection was involved. The analysis code, the derived analytic datasets, and the mapping of each variable to its source field are described in the Data Availability Statement.

2.2. Discovery Cohort

Tumors in the discovery cohort were profiled with the MSK-IMPACT targeted sequencing assay [8] across three panel versions (IMPACT341, IMPACT410, and IMPACT468). Whole-genome doubling (WGD) status was taken from the FACETS-based call provided in the dataset (field FACETS_WGD) [16]. The fraction of genome altered (FGA, range 0 to 1) was taken from the dataset field FGA_REPORTED. In the original publication, FGA was defined as the fraction of the genome with a log2 copy-number gain above 0.2 or loss below −0.2, relative to the size of the genome for which copy number was profiled [9]. We used the reported values and did not recompute them; further details of the calculation, such as segmentation and normalization, are not provided in the public files. Because the denominator is the portion of the genome covered by each targeted panel, FGA in this cohort is a panel-based estimate of genome-wide copy-number burden, and its value may be influenced by panel version, tumor purity, and sequencing coverage; panel version, panel-specific analyses, and a purity restriction were therefore examined as sensitivity analyses. Overall survival (OS) was defined as time to death from any cause, with patients alive at last contact censored. The data source describes the survival time as months since initial diagnosis, but the original publication reports survival from the time of sequencing [9], and three observations in the data are consistent with this: among censored patients, follow-up decreased with later patient accession number (Spearman correlation, −0.63); median censored follow-up differed by panel version (70.2, 52.5, and 25.9 months for IMPACT341, IMPACT410, and IMPACT468, respectively); and the maximum observed follow-up was 80.6 months, consistent with the sequencing period reported in the original publication [9]. The interval between diagnosis and sequencing is not available in the public files and could not be recovered. Age was analyzed in years at the sequencing report, with ages above 90 years set to 90.

Of 2138 sarcomas, 1774 had OS and WGD data (723 deaths) and formed the analytic set for unadjusted analyses. Subtype-stratified models were restricted to the 17 histologic subtypes represented by at least 30 patients, with complete data for FGA, age, and sample type (primary or metastasis), which left 1494 patients (593 deaths). The threshold of 30 patients was a pragmatic choice to allow stable estimation of subtype-specific baseline hazards; it is arbitrary, other thresholds were examined in sensitivity analyses, and the results apply to the more common histologies represented in this database and not to all sarcoma types. Histologic subtype was defined by the detailed cancer-type field of the dataset. Sample type describes the site of the sequenced specimen and is not a measure of disease stage; the specimen may have been obtained at any point after diagnosis. Unless

stated otherwise, all discovery models, including the extended and sensitivity models, were fitted in these 1494 patients, and the sample sizes of models fitted in other sets are given in Section 3.

2.3. Replication Cohort

Of 255 TCGA sarcomas, 252 had OS and FGA data (96 deaths); the FGA-only model, stratified by subtype and adjusted for age, included the 241 patients (93 deaths) from the five subtypes with at least 10 patients (a lower threshold than in the discovery cohort because of the smaller sample size). The TCGA cohort consisted of all PanCancer Atlas sarcoma patients with available data; it was not restricted to the pathology-reviewed cases of the original TCGA publication [6], and the diagnostic and quality-control exclusions of that publication were not applied. TCGA FGA is the Fraction Genome Altered attribute that cBioPortal derives from the segmented genome-wide copy-number data; we used it as provided and did not recompute it, and the threshold and denominator used by the portal are not reproduced in the files we analyzed. OS in this cohort is recorded by the data source as months since initial diagnosis, and age is age at diagnosis, so the survival time origin differs from that of the discovery cohort, in which follow-up starts at sequencing. WGD was derived from ABSOLUTE [17] genome-doubling calls released with the PanCancer Atlas [4], with a tumor considered WGD-positive if at least one doubling was called. Calls were available for 244 samples, all with complete OS, FGA, and age data (92 deaths); this set was used for the direct comparison of WGD and FGA, and its subtype-stratified subset comprised 233 patients (89 deaths) from the same five subtypes. All TCGA samples were primary tumors. The PanCancer Atlas clinical file used here does not contain stage, grade, tumor size, or treatment variables. Some of these annotations are available for the pathology-reviewed patients of the original TCGA sarcoma publication [6] in a separate cBioPortal study (sarc_tcga_pub), to which 206 of the 255 patients could be matched. Of the 241 patients of the FGA-only model, 200 could be matched to these annotations; among them, FNCLCC grade was recorded for 200, tumor size for 193, resection margin status for 188 (excluding 12 in whom it could not be assessed), and adjuvant radiation therapy and adjuvant systemic therapy for 191 each (excluding 9 with unknown status). We did not identify an AJCC stage field. The main TCGA models do not adjust for these annotations and include age and subtype only. In a sensitivity analysis restricted to the 193 patients with both grade and tumor size recorded, we refitted the original model (FGA and age, stratified by subtype) and then added FNCLCC grade (ordinal, 1 to 3) and tumor size (cm, linear) in the same patients, so that any change in the FGA estimate could be attributed to clinical adjustment and not to the restriction to patients with complete data.

2.4 Statistical Analysis

Associations with OS were estimated with Cox proportional hazards models using Efron's method for ties. Because a one-unit increase in FGA spans its entire possible range (0 to 1) and therefore appears large, hazard ratios (HRs) for FGA are reported per 0.10 increase, per interquartile range (IQR), and per unit increase, and the per-0.10 estimate is emphasized as the most readily interpretable. The primary discovery model was stratified by histologic subtype, giving each subtype its own baseline hazard but retaining a common coefficient for each covariate across subtypes, and included FGA, WGD, age, and sample type. To identify at which step each marker's estimate changed, WGD and FGA were each fitted in the same patients as terms were added sequentially: alone, with subtype strata, with age and sample type, and with the other marker. The FGA-OS relationship was also examined by quartiles, with cut points defined in the pooled cohort and tested as an ordinal trend; quartiles defined within each subtype served as a

sensitivity analysis. The overall difference among quartiles was tested with a likelihood-ratio test of three quartile indicators in the subtype-stratified Cox model, and the Kaplan–Meier curves by quartile are pooled and unadjusted. The log-linear form of the FGA effect was examined with restricted cubic splines (four knots at the 5th, 35th, 65th, and 95th percentiles in the discovery cohort and three knots at the 10th, 50th, and 90th percentiles in TCGA) and compared with the linear term by a likelihood-ratio test; age was entered as a linear term, and its functional form was not formally assessed.

The incremental prognostic value of each marker was assessed with likelihood-ratio tests between nested models, adding each marker to a base model of age and sample type and adding each marker to a model that already contained the other. Discrimination was summarized by Harrell's concordance index (C-index) calculated within subtype strata: concordant, discordant, and tied pairs were counted only between patients of the same subtype and pooled across subtypes, with the linear predictor of the stratified Cox model used as the risk score. Optimism was estimated from 500 bootstrap resamples of patients [18]: in each resample all models were refitted, the C-index of each resample-fitted model in the resample was compared with its C-index in the original data, and the mean difference was subtracted from the apparent C-index. Optimism-corrected C-indices and their differences between models are reported; the confidence intervals for differences between models are percentile intervals (2.5th to 97.5th percentiles) of the apparent differences across resamples and are not themselves corrected for optimism. In the TCGA cohort, the coefficients were re-estimated within that cohort, so its C-indices describe models fitted in TCGA and not the performance of a fixed model derived in the discovery cohort. An extended model additionally included sex, primary site, and panel version in the discovery cohort.

The proportional hazards assumption was evaluated with scaled Schoenfeld residuals. Because the assumption was doubtful for WGD, time-dependent effects were examined by estimating separate coefficients for the intervals 0 to 12, 12 to 36, and more than 36 months, and comparing this model with a constant-effect model by a likelihood-ratio test.

Robustness was examined in sensitivity analyses that omitted the sample-type covariate, stratified jointly by subtype and sample type, excluded gastrointestinal stromal tumors, varied the minimum subtype size (10, 20, 30, or 50 patients, or no minimum), restricted the cohort to samples with tumor purity of at least 20%, analyzed primary and metastatic samples separately, repeated the analysis within each panel version and with joint stratification by subtype and panel, tested an FGA-by-panel interaction by likelihood ratio, and omitted each subtype in turn. To quantify variation in the FGA effect across histologies, we fitted FGA-and-age models within each of the 15 subtypes with at least 10 events, pooled the estimates with a DerSimonian-Laird random-effects model [19], and report Cochran's Q and the I^2 statistic [20]. A subtype-by-FGA interaction was also tested by likelihood ratio. Subtype-specific sample sizes, event counts, FGA distributions, and hazard ratios are provided in Supplementary Table S1. To compare the FGA estimate between cohorts, the difference between the log hazard ratios of the discovery model (FGA, age, and sample type) and the TCGA model (FGA and age) was tested with a z-test.

All tests were two-sided, and p values below 0.05 were considered statistically significant. No correction for multiple testing was applied, and secondary analyses should be regarded as exploratory. Kaplan–Meier curves were compared with the log-rank test. Analyses were performed in Python (version 3.13; Python Software Foundation, Wilmington, DE, USA) using the open-source libraries lifelines (version 0.30.3) [21], pandas (version 2.3.3; NumFOCUS, Austin, TX, USA), and NumPy (version 2.5.1; NumFOCUS, Austin, TX, USA).

3. Results

3.1. Patients

Of 2138 patients in the MSK-IMPACT sarcoma cohort, 1774 had complete OS and WGD data and formed the unadjusted analytic set (723 deaths). The subtype-stratified primary analysis included 1494 patients from the 17 subtypes with at least 30 patients each (593 deaths) (Figure 1). Baseline characteristics of this primary analytic set are shown in Table 1: median age was 55 years (IQR, 38 to 66), 50.7% were male, 365 patients (24.4%) were WGD-positive, median FGA was 0.197 (IQR, 0.072 to 0.355), and 38.1% of samples were from a metastatic site. In the TCGA-SARC replication cohort, the FGA-only model included 241 patients (93 deaths; median age, 62 years [IQR, 53 to 71]; median FGA, 0.336 [IQR, 0.144 to 0.576]), and the subset with an ABSOLUTE-based WGD call comprised 244 patients (92 deaths), of whom 115 (47.1%) were WGD-positive. All TCGA samples were primary tumors.

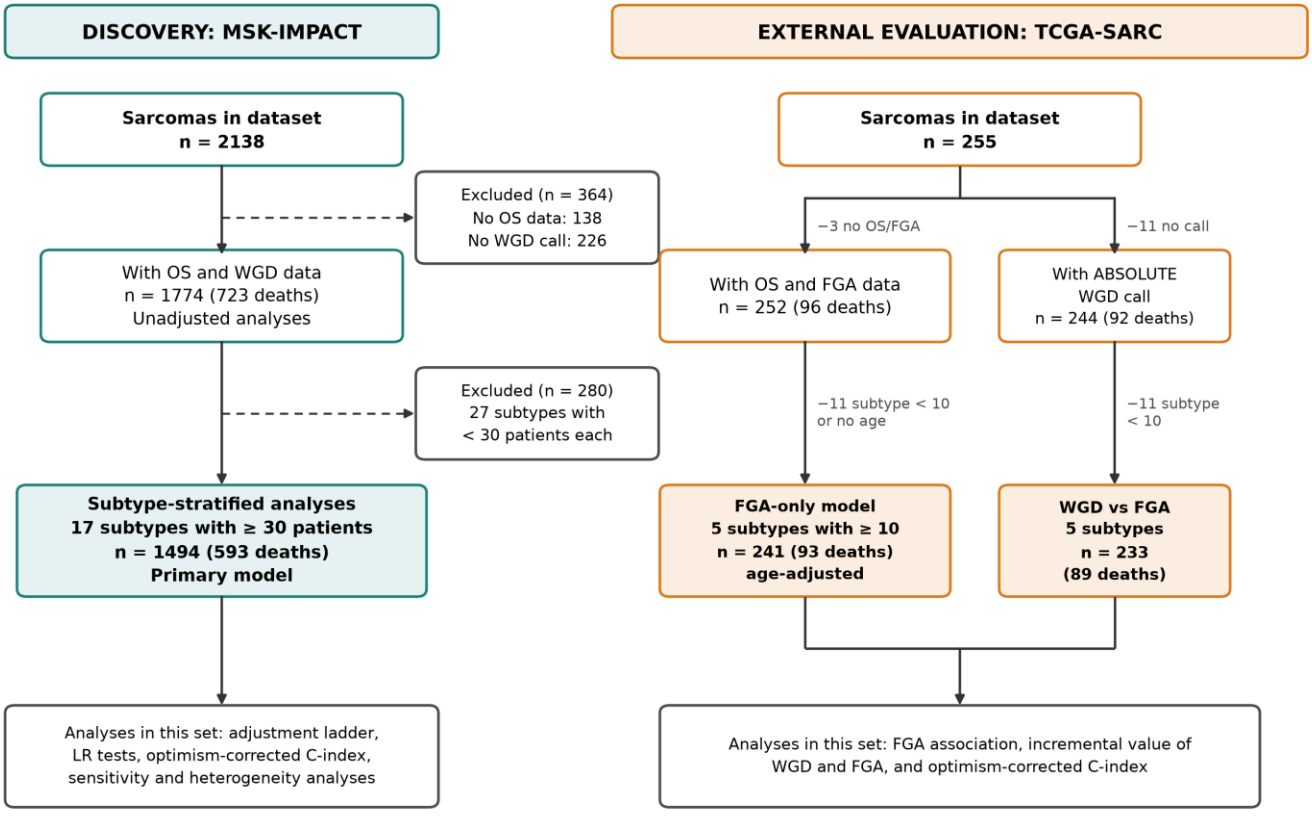


Figure 1. Patient flow and analytic cohorts in the MSK-IMPACT discovery cohort and the TCGA-SARC external replication cohort.

Table 1. Baseline characteristics of the discovery and TCGA-SARC analytic sets.

Characteristic	Discovery, Unadjusted Set (n = 1774)	Discovery, Primary Model (n = 1494)	TCGA-SARC, FGA-Only Model (n = 241)	TCGA-SARC, WGD-FGA Comparison Set (n = 244)
Deaths, n (%)	723 (40.8)	593 (39.7)	93 (38.6)	92 (37.7)
Age, years, median (IQR)	54 (36–65)	55 (38–66)	62 (53–71)	61 (53–71)
Male sex, %	48.4	50.7	46.1	45.9
WGD-positive, n (%)	407 (22.9)	365 (24.4)	see next column †	115 (47.1) *
FGA, median (IQR)	0.184 (0.060–0.350)	0.197 (0.072–0.355)	0.336 (0.144–0.576)	0.340 (0.147–0.583)
Metastatic sample, %	38.5	38.1	0 (all primary)	0 (all primary)
Histologic subtypes included	44	17 (≥30 patients each)	5 (≥10 patients each)	7 (5 with ≥10 patients)

IQR, interquartile range; WGD, whole-genome doubling; FGA, fraction of genome altered. * WGD was FACETS-based in the discovery cohort and ABSOLUTE-based in TCGA-SARC (WGD-positive: at least one genome doubling called). † The 244 patients with an ABSOLUTE-based WGD call overlap with the 241 patients of the FGA-only model; 233 of them, from the five subtypes with at least 10 patients, were used in the stratified WGD-versus-FGA comparison (see Section 3.8).

WGD prevalence varied widely across the 17 discovery subtypes, from 5.0% in myxoid/round-cell liposarcoma to 51.9% in osteosarcoma (Figure 2), which is consistent with the possibility that an unadjusted association between WGD and survival partly reflects differences in histologic composition rather than a within-subtype effect.

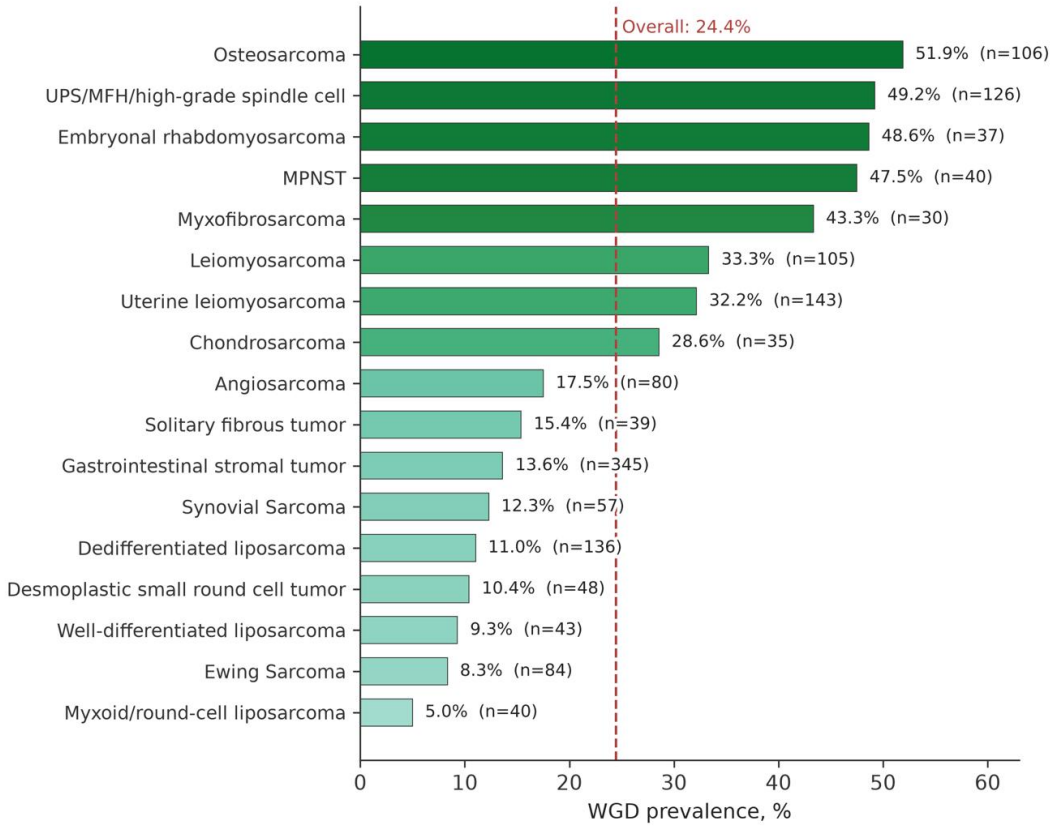
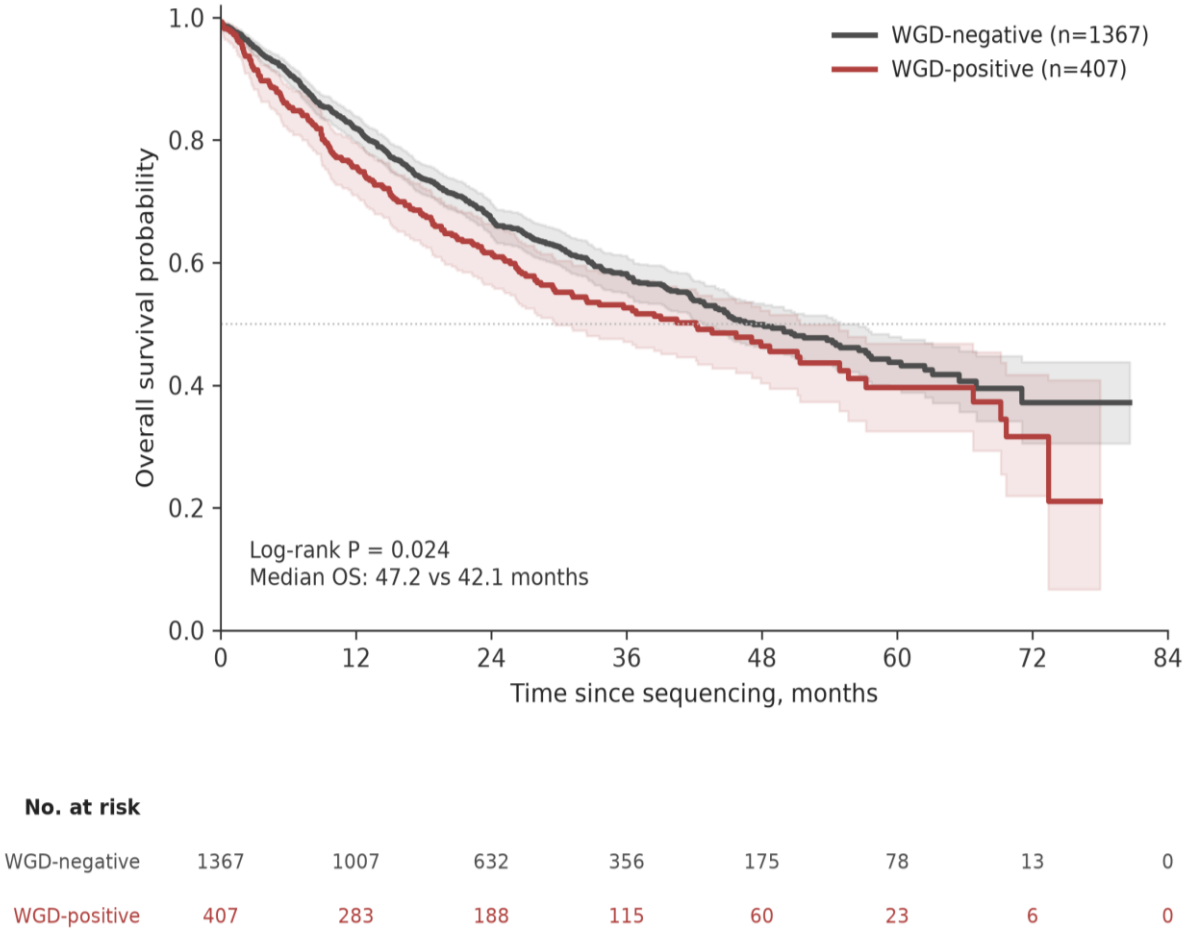


Figure 2. Prevalence of whole-genome doubling (WGD) across the 17 histologic subtypes of the primary analytic set (n = 1494). Bars show the percentage of WGD-positive tumors (FACETS-based call) in each subtype, with the number of patients in parentheses; the dashed line marks the overall prevalence.

3.2. Unadjusted Associations and the Effect of Subtype Adjustment

In the unadjusted analytic set (n = 1774), both markers were associated with shorter OS: WGD (HR, 1.21; 95% CI, 1.02 to 1.43; $p = 0.024$) and FGA (HR per unit increase, 2.69; 95% CI, 1.92 to 3.76; $p < 0.001$). Median OS was 42.1 months for WGD-positive patients versus 47.2 months for WGD-negative patients (log-rank $p = 0.024$) (Figure 3).



No. at risk								
WGD-negative	1367	1007	632	356	175	78	13	0
WGD-positive	407	283	188	115	60	23	6	0

Figure 3. Overall survival by whole-genome doubling (WGD) status in the unadjusted analytic set of the discovery cohort (n = 1774; 723 deaths). Kaplan–Meier curves are unadjusted for subtype, age, and sample type; the *p* value is from a log-rank test.

To determine whether these associations reflected a within-subtype effect, we re-fit each marker in the same 1494 patients as covariates were added sequentially (Figure 4). In contrast to the log-rank comparison above, which used the 1774-patient analytic set, all estimates in Figure 4 come from Cox models in the 1494 patients retained for the stratified analysis, which is why the unadjusted WGD hazard ratio is 1.25 here and 1.21 above. For WGD, the hazard ratio decreased and lost statistical significance as soon as subtype strata were introduced (unadjusted, 1.25 [95% CI, 1.05 to 1.50]; plus subtype strata, 1.12 [0.93 to 1.36]; plus age and sample type, 1.14 [0.94 to 1.38]) and moved to the null once FGA was added (0.98 [0.80 to 1.19]; $p = 0.81$). For FGA, the hazard ratio was stable across the same sequence of adjustments (unadjusted, 2.71 [1.87 to 3.93]; plus subtype strata, 2.69 [1.75 to 4.13]; plus age and sample type, 2.79 [1.82 to 4.26]; plus WGD, 2.84 [1.81 to 4.46]; $p < 0.001$ throughout). FGA and WGD were moderately correlated (Spearman $\rho = 0.44$).

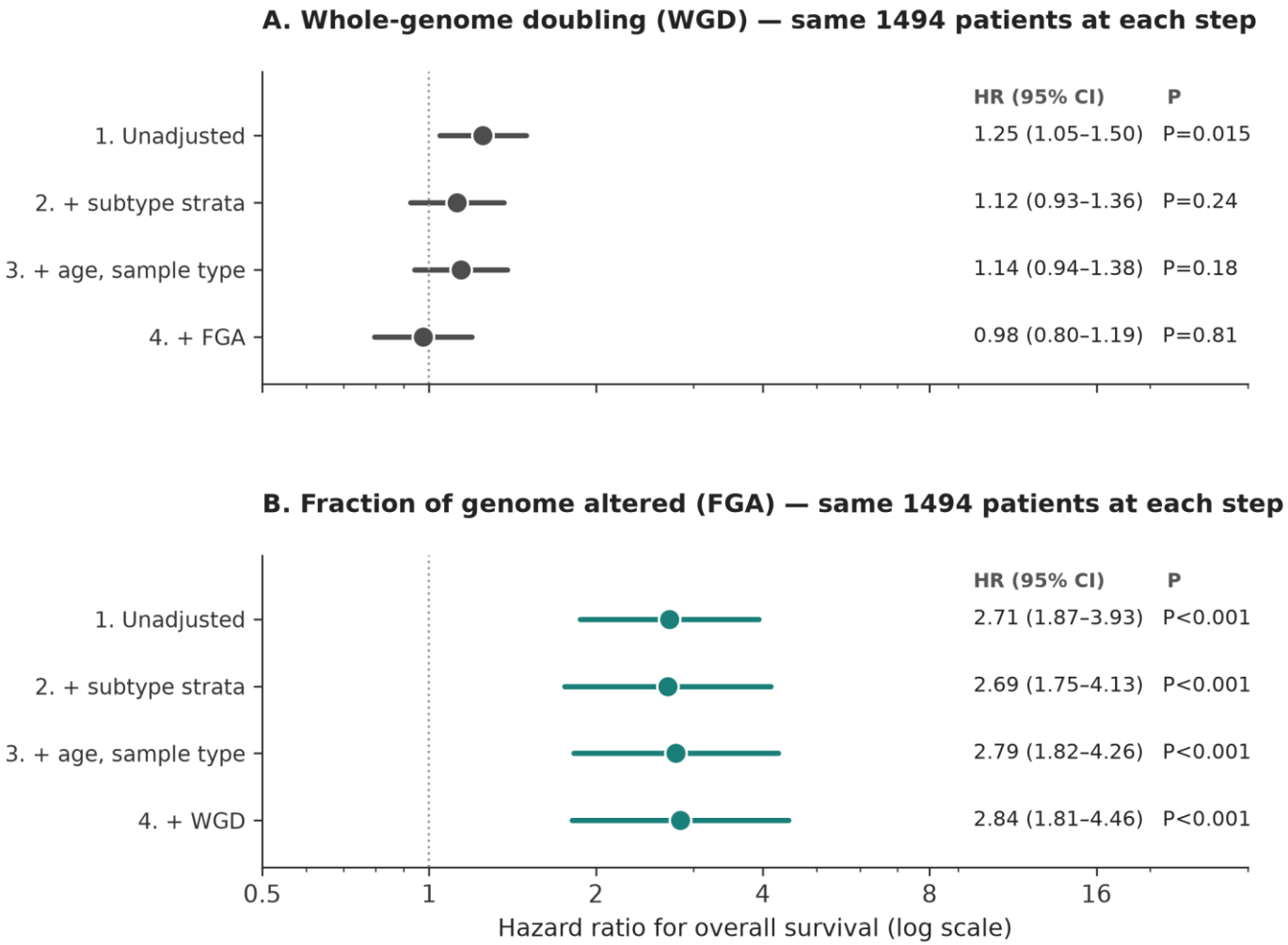


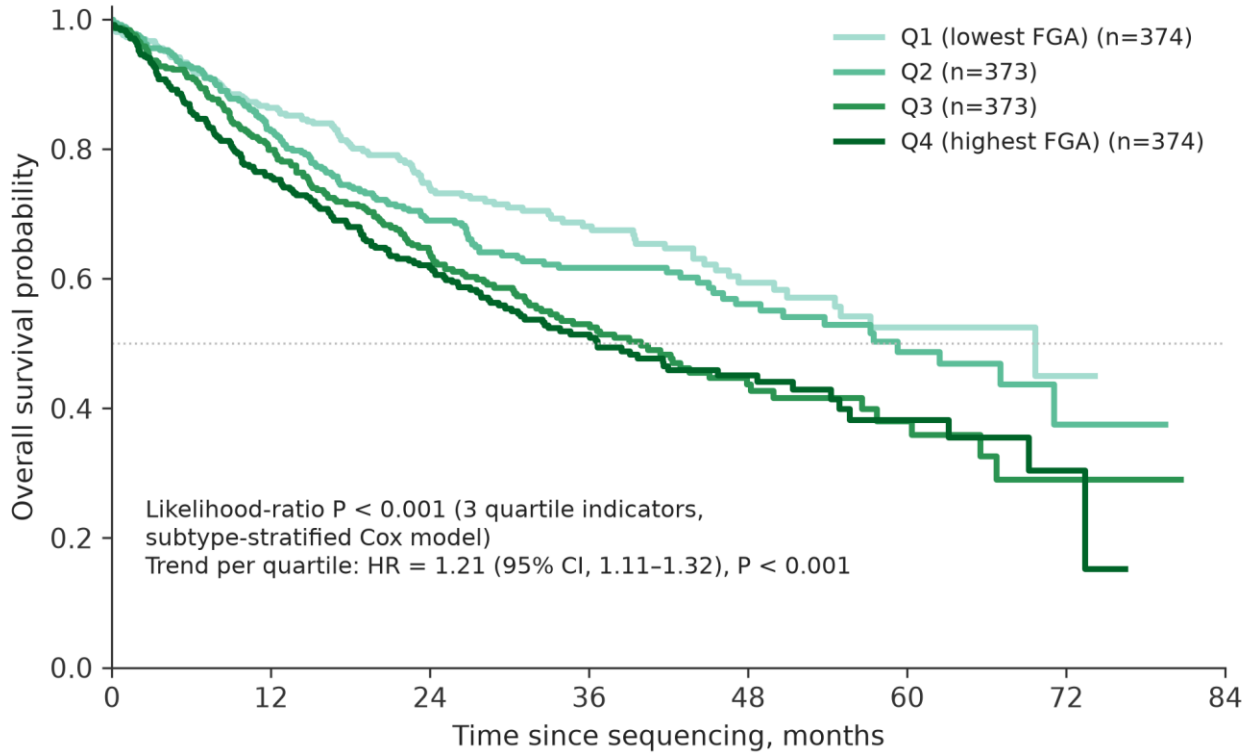
Figure 4. Effect of sequential covariate adjustment on the hazard ratios for WGD and FGA. Hazard ratios (95% CI) for WGD (**A**) and for FGA per unit increase (**B**) are from Cox models fitted in the same 1494 patients at each step: unadjusted, with subtype strata, with age and sample type added, and with the other marker added. *p* values are from Wald tests of the coefficient in each model; they differ from the log-rank *p* value of Figure 3, which was derived from the 1774-patient analytic set.

3.3. Primary Model

In the fully specified subtype-stratified model (n = 1494; 593 deaths), which included FGA, WGD, age, and sample type, FGA remained associated with OS (HR per 0.10 increase, 1.11; 95% CI, 1.06 to 1.16; HR per interquartile range, 1.34; 95% CI, 1.18 to 1.53; HR per unit increase, 2.84; 95% CI, 1.81 to 4.46; $p < 0.001$). WGD was not associated with OS in this model (HR, 0.98; 95% CI, 0.80 to 1.19; $p = 0.81$). Older age (HR, 1.01 per year; 95% CI, 1.01 to 1.02; $p < 0.001$) and a metastatic sample (HR, 1.79; 95% CI, 1.51 to 2.12; $p < 0.001$) were associated with shorter OS in the same model.

The relationship between FGA and OS was graded rather than confined to a single threshold. Across FGA quartiles defined in the pooled cohort, hazard ratios relative to the lowest quartile increased from 1.28 (95% CI, 0.99 to 1.66) to 1.53 (1.18 to 1.98) to 1.80 (1.38 to 2.36) in the subtype-stratified model, with a significant ordinal trend (HR per quartile step, 1.21; 95% CI, 1.11 to 1.32; $p < 0.001$) (Figure 5). The pooled quartile boundaries were FGA values of 0.072, 0.197, and 0.355, and a likelihood-ratio test of the three quartile indicators in the subtype-stratified model was significant ($\chi^2 = 21.1$; df = 3; $p < 0.001$). Quartiles defined within each subtype showed a similar trend (HR per step, 1.22; 95% CI,

1.13 to 1.31; $p < 0.001$). On the log-hazard scale, restricted cubic splines showed no significant departure from a linear FGA effect in either cohort ($p = 0.15$ in the discovery cohort and $p = 0.70$ in TCGA), which supports the log-linear form used for the per-0.10 and per-unit hazard ratios, although these tests have limited power.



No. at risk

Q1	374	287	184	110	58	25	1	0
Q2	373	277	185	96	62	29	5	0
Q3	373	274	172	101	44	18	4	0
Q4	374	259	170	102	48	19	4	0

Figure 5. Overall survival by quartile of FGA in the primary analytic set of the discovery cohort (n = 1494). Quartile boundaries were defined in the pooled cohort (FGA 0.072, 0.197, and 0.355). Kaplan–Meier curves are pooled and unadjusted; the *p* value is from a likelihood-ratio test of the three quartile indicators in a Cox model stratified by subtype, and not from a log-rank test.

3.4. Incremental Prognostic Value

Likelihood-ratio tests in the subtype-stratified model (age and sample type as the base model) showed that adding FGA improved model fit ($\chi^2 = 21.4$; $p < 0.001$), whereas adding WGD did not ($\chi^2 = 1.8$; $p = 0.18$). When each marker was added to a model already containing the other, FGA still improved fit over WGD ($\chi^2 = 19.7$; $p < 0.001$), but WGD added nothing over FGA ($\chi^2 = 0.06$; $p = 0.81$) (Table 2). This pattern was unchanged in an extended model that also adjusted for sex, primary tumor site, and sequencing panel (FGA|WGD, $p < 0.001$; WGD|FGA, $p = 0.89$). Because FGA and WGD were correlated (Spearman $\varrho = 0.44$; median FGA 0.14 in WGD-negative and 0.42 in WGD-positive tumors in the unadjusted analytic set; Figure 6), these nested comparisons describe the

information that each marker carries conditional on the other, and not the performance of either marker alone.

Table 2. Incremental prognostic value of FGA and WGD.

Comparison	Cohort	χ^2	df	*p*
WGD added to age + sample type	Discovery (n = 1494)	1.8	1	0.18
FGA added to age + sample type	Discovery (n = 1494)	21.4	1	<0.001
FGA added to WGD + age + sample type	Discovery (n = 1494)	19.7	1	<0.001
WGD added to FGA + age + sample type	Discovery (n = 1494)	0.06	1	0.81
WGD added to age	TCGA-SARC (n = 233)	4.9	1	0.03
FGA added to age	TCGA-SARC (n = 233)	12.8	1	<0.001
FGA added to WGD + age	TCGA-SARC (n = 233)	8.5	1	0.004
WGD added to FGA + age	TCGA-SARC (n = 233)	0.6	1	0.44

All comparisons are likelihood-ratio tests from subtype-stratified Cox models. FGA, fraction of genome altered; WGD, whole-genome doubling. Discovery models include age and sample type as covariates and TCGA-SARC models include age; TCGA-SARC coefficients were estimated within that cohort.

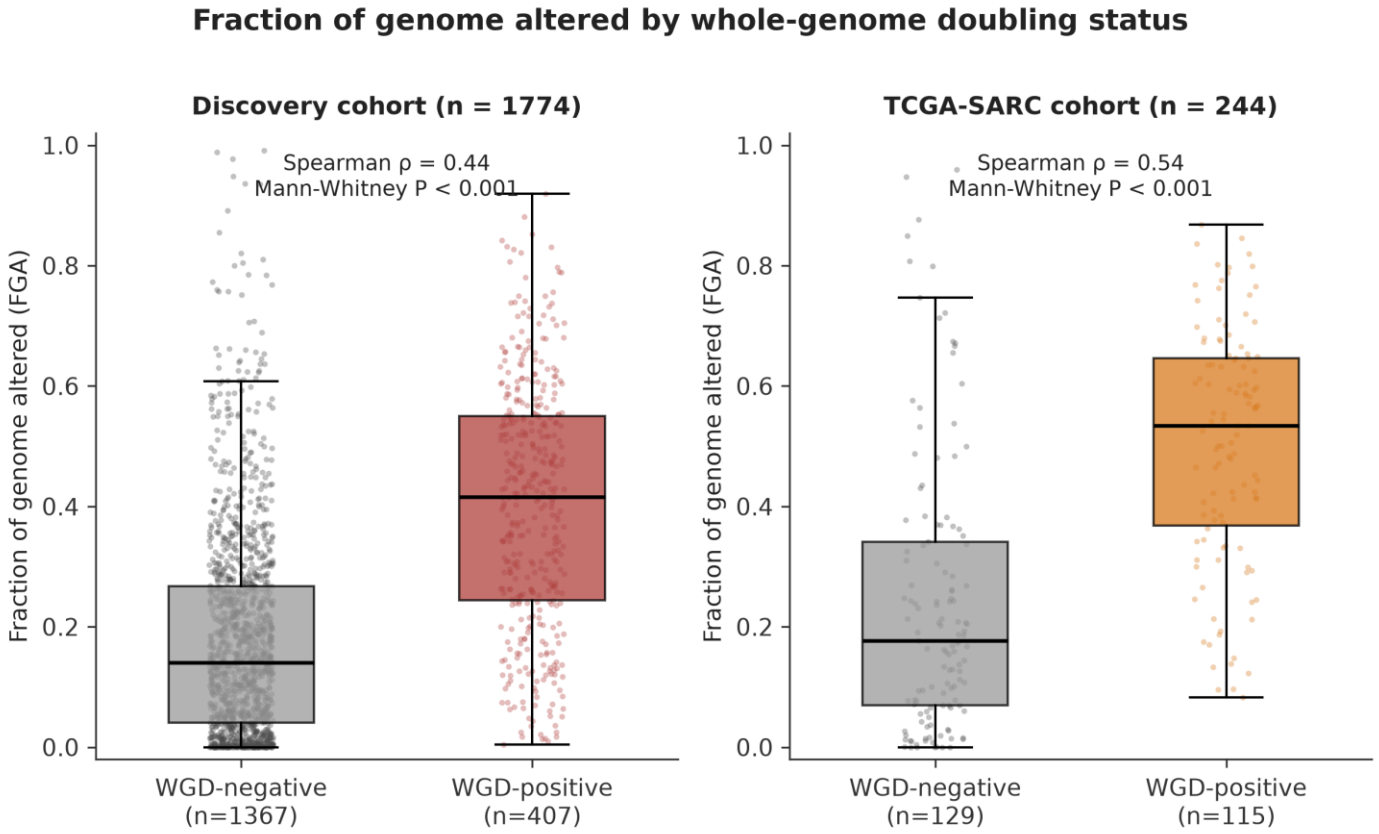


Figure 6. Fraction of genome altered (FGA) by whole-genome doubling (WGD) status in the unadjusted analytic set of the discovery cohort (left; n = 1774; FACETS-based WGD) and in TCGA-SARC patients with an ABSOLUTE-based WGD call (right; n = 244). Boxes show medians and interquartile ranges, and points are individual tumors. FGA was higher in WGD-positive tumors in both cohorts, and the two measures were correlated (Spearman ρ = 0.44 and 0.54, respectively).

Optimism-corrected C-indices (500 bootstrap resamples) showed a small but consistent gain from adding FGA to the clinical model (age and sample type): 0.618 for the clinical model, 0.618 with WGD added, and 0.633 with FGA added (both markers,

0.632). The corrected difference in C-index between the FGA model and the WGD model was 0.015 (bootstrap 95% CI, 0.001 to 0.032); adding WGD on top of FGA gave essentially no further gain (corrected difference, −0.001; 95% CI, −0.002 to 0.007). The extended model showed a similar pattern (corrected ΔC for FGA versus WGD, 0.011; 95% CI, −0.001 to 0.022). The model containing both markers had a marginally lower corrected C-index than the model with FGA alone (0.632 versus 0.633), which is consistent with an uninformative additional term adding estimation noise and optimism rather than with an inconsistency in the results. Here and for the TCGA cohort below, the point estimates of the C-index differences are optimism-corrected, whereas the confidence intervals reported for these differences are percentile intervals of the uncorrected (apparent) differences across the bootstrap resamples.

3.5. Proportional Hazards and Time-Varying Effects

Scaled Schoenfeld residuals showed no evidence of non-proportionality for FGA, age, or sample type (all $p > 0.1$), but WGD showed a significant rank-based violation ($p = 0.0056$). When the WGD effect was allowed to vary across three time windows (0 to 12, 12 to 36, and more than 36 months), it was significant only in the first year (HR, 1.41; 95% CI, 1.06 to 1.87) and null thereafter; a likelihood-ratio test did not favor the time-varying model over a constant-effect model ($p = 0.14$), and after adjustment for FGA, WGD was not significant in any window. FGA showed no evidence of a time-varying effect (likelihood-ratio $p = 0.58$).

3.6. Sensitivity Analyses

The association between FGA and OS was robust across a range of alternative specifications. It persisted after excluding the sample-type covariate (HR, 2.77), after jointly stratifying by subtype and sample type (HR, 2.63), and after restricting the cohort to samples with tumor purity of at least 20% (HR, 3.05). Because gastrointestinal stromal tumor had the largest single-subtype effect, we repeated the primary model excluding it (n = 1149; 522 deaths); FGA remained associated with OS (HR, 2.06; 95% CI, 1.26 to 3.36; $p = 0.004$), while WGD did not (HR, 1.02; $p = 0.89$). Varying the minimum subtype size from 10 to 50 patients, or removing the size restriction altogether, did not change the qualitative pattern (FGA HR ranging from 2.64 to 3.09 across thresholds; WGD non-significant throughout). Omitting each subtype in turn changed the FGA hazard ratio only within the range of 2.06 to 3.12, with gastrointestinal stromal tumor again the most influential subtype.

Because FGA is estimated from panels of different sizes, we repeated the analysis by panel version. The hazard ratio for FGA per unit increase was 6.27 (95% CI, 2.04 to 19.26) in IMPACT341 (n = 142; 99 deaths), 2.80 (95% CI, 1.19 to 6.58) in IMPACT410 (n = 402; 208 deaths), and 3.05 (95% CI, 1.57 to 5.93) in IMPACT468 (n = 950; 286 deaths), and WGD was not associated with OS in any panel ($p = 0.74$, 0.69, and 0.30, respectively). Stratifying jointly by subtype and panel version gave an FGA hazard ratio of 3.15 (95% CI, 1.97 to 5.02; $p < 0.001$; WGD, HR, 0.95; $p = 0.62$), and restricting the cohort to the two larger panels gave 2.58 (95% CI, 1.55 to 4.28; n = 1352; 494 deaths). The FGA-by-panel interaction was significant across the three panels (likelihood-ratio $\chi^2 = 15.7$; df = 2; $p < 0.001$) but not when IMPACT341 was excluded (IMPACT410 versus IMPACT468: $\chi^2 = 1.3$; df = 1; $p = 0.25$), indicating that the estimate was larger and less precise in the smallest and earliest panel and similar in the two later ones. When primary and metastatic samples were analyzed separately, FGA was associated with OS in both (primary samples: HR, 2.90; 95% CI, 1.50 to 5.59; n = 925; 293 deaths; metastatic samples: HR, 2.39; 95% CI, 1.24 to 4.62; n = 569; 300 deaths), whereas WGD was not (primary samples: HR, 0.90; 95% CI, 0.66 to 1.22; metastatic samples: HR, 1.07; 95% CI, 0.81 to 1.43). The extended model, the model

without the sample-type covariate, and the model stratified jointly by subtype and sample type used the same 1494 patients; the purity-restricted model used 1463 patients (572 deaths); and the models with a minimum subtype size of 10, 20, or 50 patients, or no minimum, used 1709 (691 deaths), 1589 (643 deaths), 1182 (471 deaths), and 1774 patients (723 deaths), respectively.

3.7. Heterogeneity Across Histologic Subtypes

The magnitude of the FGA effect varied across the 15 subtypes with at least 10 events (Cochran Q = 32.4; df = 14; $p = 0.004$; $I^2 = 57\%$). A random-effects pooled estimate remained consistent with the primary model (HR, 2.97; 95% CI, 1.45 to 6.09), and the hazard ratio exceeded 1 in 11 of the 15 subtypes, including gastrointestinal stromal tumor, solitary fibrous tumor, Ewing sarcoma, and synovial sarcoma. The 95% confidence interval excluded 1 in only four of these subtypes (gastrointestinal stromal tumor, Ewing sarcoma, synovial sarcoma, and solitary fibrous tumor), and the intervals were wide in most of the smaller subtypes. Subtype-specific sample sizes, event counts, FGA distributions, and hazard ratios per 0.10 increase in FGA with confidence intervals are provided in Supplementary Table S1; two subtypes with fewer than 10 events (myxoid/round-cell liposarcoma and embryonal rhabdomyosarcoma) are listed without an estimate. Because the primary model retains a single common FGA coefficient, its estimate should be read as an average across subtypes and not as evidence of an association within each histology. A likelihood-ratio test for a FGA-by-subtype interaction was not statistically significant at the 0.05 level ($\chi^2 = 25.7$; df = 16; $p = 0.058$), which does not exclude subtype-dependent variation in effect size but does not establish it as a formal statistical interaction either.

3.8. External Replication in TCGA-SARC

In the TCGA-SARC cohort, the pre-specified FGA-only model (n = 241; 93 deaths; age-adjusted, subtype-stratified) showed a hazard ratio of 5.15 (95% CI, 2.05 to 12.93; $p < 0.001$) per unit increase in FGA, in the same direction as, and numerically larger than, the discovery estimate. The interval was wide, and the TCGA estimate did not differ significantly from the discovery estimate (HR, 2.79; model with FGA, age, and sample type; difference in log hazard ratios, z = 1.19; $p = 0.24$), so these data do not show a true difference in effect size between the cohorts. In the subset with an ABSOLUTE-based WGD call (n = 233; 89 deaths), FGA remained associated with OS when WGD was included in the same model (HR, 4.74; 95% CI, 1.70 to 13.22; $p = 0.003$), whereas WGD was not (HR, 1.22; 95% CI, 0.74 to 2.00; $p = 0.44$). Likelihood-ratio tests reproduced the discovery pattern: adding FGA to a model containing WGD improved fit ($\chi^2 = 8.5$; $p = 0.004$), while adding WGD to a model containing FGA did not ($\chi^2 = 0.6$; $p = 0.44$). Unlike in the discovery cohort, WGD remained associated with OS in TCGA in a model that contained WGD and age, stratified by subtype but without FGA (HR, 1.69; 95% CI, 1.06 to 2.68; $p = 0.028$); what was reproduced across cohorts was not the loss of a WGD association per se, but the absence of detectable incremental information from WGD once FGA was known. Optimism-corrected C-indices in this cohort were 0.562 (age only), 0.578 (plus WGD), and 0.639 (plus FGA) (both markers, 0.632), with a corrected difference of 0.062 (95% CI, 0.001 to 0.115) favoring FGA over WGD. These C-indices come from models whose coefficients were estimated within TCGA and describe those models, not the performance of a fixed model derived in the discovery cohort, and the confidence interval of their difference is wide.

To examine whether the TCGA association could be explained by tumor grade and size, we performed a sensitivity analysis in the same patients. In the 193 patients with both FNCLCC grade and tumor size recorded (72 deaths), the FGA hazard ratio per unit increase was 3.91 (95% CI, 1.37 to 11.17) in the original model (FGA and age, stratified by

subtype), 2.88 (95% CI, 0.94 to 8.86) after adding grade, and 2.51 (95% CI, 0.81 to 7.72) after adding grade and tumor size; per 0.10 increase, the corresponding hazard ratios were 1.15 (95% CI, 1.03 to 1.27), 1.11 (95% CI, 0.99 to 1.24), and 1.10 (95% CI, 0.98 to 1.23). The estimate of the original model was lower in this subset than in the full cohort (3.91 versus 5.15), so part of the difference from the full-cohort estimate reflects the restriction to patients with complete data, whereas the further reduction to 2.51 reflects clinical adjustment. Grade and tumor size together improved model fit (likelihood-ratio $\chi^2 = 9.5$; df = 2; $p = 0.009$); tumor size (HR per cm, 1.04; $p = 0.009$) but not grade (HR per grade step, 1.52; $p = 0.12$) was individually associated with OS. The point estimate for FGA remained above 1, but the confidence intervals were wide and included 1 after adjustment, so these data neither exclude partial confounding by grade and tumor size nor show that the association disappears. In the 200 patients with grade recorded (76 deaths), the FGA hazard ratio was 4.53 (95% CI, 1.64 to 12.52) before and 3.36 (95% CI, 1.14 to 9.94) after adding grade, and entering tumor size as its logarithm in the 193-patient set gave a similar result (HR, 2.48; 95% CI, 0.80 to 7.67).

4. Discussion

In two independent sarcoma cohorts, higher fraction of genome altered (FGA) was associated with shorter overall survival after adjustment for histologic subtype and the available covariates (age and sample type in the discovery cohort, age in TCGA), whereas whole-genome doubling (WGD) added no detectable prognostic information once FGA was known. The two markers are related, since both are derived from the same copy-number profile, but they behaved differently once histology was taken into account: the WGD hazard ratio moved toward the null as subtype strata were introduced, while the FGA hazard ratio remained essentially unchanged at every step. This pattern suggests that the burden of copy-number alteration may carry more of the prognostic information available from routine copy-number data in this setting than doubling status does, although the two measures are correlated and this analysis cannot separate their underlying biology.

This reading is compatible with earlier work rather than opposed to it. Genomic instability has long been considered an enabling characteristic of cancer [1], and experimental studies indicate that WGD tolerates and propagates further chromosomal instability [2]. In a pan-cancer cohort, WGD was associated with worse outcomes independently of established clinical factors [3], and we do not dispute that finding at the pan-cancer level. Sarcomas, however, comprise many histologic subtypes with markedly different biology [5], and WGD prevalence varied more than ten-fold across the subtypes in our discovery cohort. A marker whose frequency tracks histology can appear prognostic in a mixed-histology analysis even if it carries little independent information once subtype is accounted for. FGA and WGD are related but distinct measures: FGA quantifies the extent of the genome affected by copy-number alteration, whereas WGD is an inferred genome-doubling event, and our results compare the conditional prognostic information carried by the two measures without showing that any difference between them arises from continuous versus binary measurement alone. A related secondary analysis of the same discovery cohort combined WGD and FGA with four other genomic features into a single composite instability score that was also prognostic [10]; because that analysis combined markers rather than comparing them, it could not establish whether WGD adds information beyond FGA, which was the specific question addressed here, together with external replication in an independent cohort.

The gain in discrimination from FGA was modest. The optimism-corrected C-index increased by 0.01 to 0.06 across our models, which indicates a small improvement over a limited clinical comparator (age and sample type) and does not establish a clinically useful

improvement over a comprehensive prognostic assessment that includes grade, stage, and treatment. The TCGA analysis should be regarded as an external replication of the association between FGA and survival, and of the pattern of incremental information, rather than as an external validation of a prediction model, because the coefficients were re-estimated in that cohort and the confidence intervals were wide. In practical terms, this study speaks less to whether genomic instability should be measured at all, since FGA is already calculated from the copy-number data generated by clinical next-generation sequencing panels [8], and more to which of two already-available measures deserves more attention when, in the evaluated models, only one adds detectable incremental information.

The time origin has direct consequences for how the primary result should be interpreted. In the discovery cohort, follow-up begins at sequencing, so each patient had to survive from diagnosis until a tumor sample was sequenced, and both FGA and WGD were measured at that time point. The hazard ratios therefore describe the association between copy-number burden at sequencing and subsequent survival, and not between copy-number burden at diagnosis and survival from diagnosis. The direction of any resulting bias cannot be established from these data. One plausible mechanism is that patients with the most aggressive disease, who may be more likely to have high FGA, die before sequencing and are therefore under-represented, which would tend to attenuate the observed association; sequencing later in the disease course, for example from a metastatic site, could act in either direction. On balance we consider attenuation more likely than inflation, but this is an expectation and not a finding. Because the interval between diagnosis and sequencing is not recorded in the public dataset, a delayed-entry (left-truncation) analysis could not be performed. In TCGA, survival is recorded from initial diagnosis and all tumors are primary samples, so the two cohorts differ in time origin as well as in platform, and their hazard ratios should not be treated as estimates of the same quantity. We have accordingly avoided interpreting our estimates as diagnosis-anchored prognostic effects.

A further consideration is that FGA may partly reflect tumor grade or other aggressive tumor features, and grade is among the most established prognostic factors in soft-tissue sarcoma. Neither grade nor stage was available in the discovery cohort, and in TCGA these annotations are not part of the PanCancer Atlas clinical file that we used. The comparison made here is therefore between two genomics-derived measures, FGA and WGD, and not between FGA and standard clinical prognostication, and the incremental value of FGA over grade, stage, tumor size, resection margin, or treatment is not established in this study beyond the limited TCGA sensitivity analysis described below. For this reason we describe our findings as associations conditional on the covariates listed for each model and do not describe FGA as an independent prognostic factor. Grade and tumor size are recorded for most TCGA patients in the cBioPortal study of the original TCGA publication, as noted in the Methods. In a sensitivity analysis in the 193 patients with both variables recorded, the FGA hazard ratio decreased from 3.91 to 2.51 after adding them, and its confidence interval, which was wide because of the small number of deaths (72), then included 1; the association may therefore be partly explained by grade and tumor size, and this analysis cannot establish whether FGA carries prognostic information beyond them.

FGA was also not measured in the same way across panels or cohorts. In the discovery cohort it is a panel-based approximation of genome-wide copy-number burden, whereas in TCGA it is derived from genome-wide data. The higher hazard ratio for FGA in TCGA (5.15 versus 2.84) should therefore not be read as a true difference in effect between cohorts: the TCGA interval is wide, the difference between the two log hazard ratios was not statistically significant ($p = 0.24$), and the cohorts differ in the measurement

of FGA, the time origin, the number of subtypes, and the selection of patients. Within the discovery cohort, the FGA estimate differed by panel version (6.27 for IMPACT341, 2.80 for IMPACT410, and 3.05 for IMPACT468), with a significant interaction overall that was driven by the smallest and earliest panel; panel version is correlated with accession order and follow-up, which links this issue to the time-origin considerations above.

Several limitations warrant emphasis. First, overall survival in the discovery cohort was measured from the time of sequencing rather than from diagnosis, which restricts the cohort to patients who survived long enough to be sequenced and limits the interpretation of our estimates as diagnosis-anchored prognostic effects, as discussed above; the interval between diagnosis and sequencing could not be recovered, so a delayed-entry analysis was not possible. Second, stage, grade, tumor size, and treatment were not available in the discovery cohort, and although grade and tumor size are recorded for most TCGA patients in a separate cBioPortal study and were examined only in a small sensitivity analysis (193 patients, 72 deaths) with wide confidence intervals; residual confounding by unmeasured clinical factors, including the possibility that FGA is partly a marker of grade, therefore cannot be excluded, and sample type is a measure of specimen origin and not of disease stage. Third, FGA was derived from targeted panels of different sizes in the discovery cohort and from genome-wide data in TCGA, and WGD was called with different algorithms, FACETS in the discovery cohort and ABSOLUTE in TCGA-SARC; although both estimate genome-doubling status from copy-number data, the methods are not identical, which limits the precision of any direct cross-cohort comparison. Fourth, the magnitude of the FGA effect varied across histologic subtypes ($I^2 = 57\%$) and across panel versions, and the primary models estimate a single common FGA coefficient, so the estimates are an average across a heterogeneous group of diseases, and a significant pooled coefficient does not imply an association in every histology; the analysis is also limited to subtypes with at least 30 patients (10 in TCGA) and applies to the more common histologies in these databases. Fifth, the proportional hazards assumption was doubtful for WGD, and although the time-varying analyses did not change the conclusion, a single hazard ratio may not fully describe its effect; in addition, the log-linear form of age was not formally assessed. Finally, our sensitivity analyses were exploratory and not adjusted for multiple comparisons, and both cohorts were assembled retrospectively from clinically sequenced patients rather than from a population-based inception cohort.

5. Conclusions

A higher fraction of genome altered was associated with shorter overall survival in two sarcoma cohorts profiled on different sequencing platforms, after adjustment for histologic subtype, age, and, in the discovery cohort, sample type, whereas whole-genome doubling added no detectable incremental information once FGA was known in the evaluated models; these findings are subject to uncertainty and may differ within particular subtypes. Because FGA is already generated by the copy-number analysis underlying routine clinical sequencing, it may warrant more attention than WGD status when genomic complexity is considered alongside histology in the interpretation of a sarcoma genomic report. The gain in discrimination was modest; however, survival was measured from sequencing in the discovery cohort, and stage, grade, and treatment were not included in the main analyses; prospective studies that incorporate these clinical factors will be needed before genomic instability measures can inform treatment decisions in sarcoma.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/diagnostics16233974/s1, Table S1: Subtype-specific results in the discovery cohort (primary analytic set, n = 1494). For each subtype: number of patients, deaths, WGD prevalence, FGA median (interquartile range), and the hazard ratio per 0.10 increase in FGA (95% CI) from a Cox model fitted within that subtype with FGA and age as covariates (the within-subtype models used for the heterogeneity analysis; sample type and WGD not included). Models were fitted only in subtypes with at least 10 deaths. The primary model, by contrast, retains a single common FGA coefficient across subtypes. Cochran Q = 32.4 (df = 14; p = .004); I^2 = 57%; random-effects pooled hazard ratio per unit FGA 2.97 (95% CI, 1.45–6.09). Table S2: Sample sizes and numbers of deaths of all analytic sets and models reported in the manuscript. Table S3: Mapping of analysis variables to source fields. Table S4. TCGA-SARC cohort construction and completeness of clinical annotations in the separate cBioPortal study of the original TCGA publication (sarc_tcga_pub). These annotations were used only in the sensitivity analysis restricted to patients with grade and tumor size recorded; they were not used in any other model.

Author Contributions: Conceptualization, M.M.Ç.; methodology, M.M.Ç.; software, M.M.Ç.; validation, M.M.Ç. and M.A.; formal analysis, M.M.Ç.; investigation, M.M.Ç.; resources, M.M.Ç.; data curation, M.M.Ç.; writing—original draft preparation, M.M.Ç. and M.A.; writing—review and editing, M.M.Ç. and M.A.; visualization, M.M.Ç.; supervision, M.M.Ç.; project administration, M.M.Ç. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study was a secondary analysis of publicly available, de-identified data obtained from cBioPortal (study identifiers sarcoma_mskcc_2022 and sarc_tcga_pan_can_atlas_2018). No new human subjects were recruited, and no new clinical or biological data were collected; ethical review and institutional approval were therefore not required for this analysis.

Informed Consent Statement: Not applicable. This study did not involve the recruitment of new human subjects. All data used were obtained from publicly available, de-identified repositories, in which informed consent was obtained by the original data-generating studies at the time of collection.

Data Availability Statement: The source data analyzed in this study are publicly available from cBioPortal (https://www.cbioportal.org) under the study identifiers sarcoma_mskcc_2022 (MSK-IMPACT sarcoma cohort) and sarc_tcga_pan_can_atlas_2018 (TCGA-SARC, PanCancer Atlas); the ABSOLUTE genome-doubling calls were obtained from the PanCancer Atlas resources of the Genomic Data Commons. The analysis code, the derived analytic datasets, and a table mapping each variable to its source The analysis code, the derived analytic datasets, and a table mapping each variable to its source field are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674.
2. Dewhurst, S.M.; McGranahan, N.; Burrell, R.A.; Rowan, A.J.; Grönroos, E.; Endesfelder, D.; Joshi, T.; Mouradov, D.; Gibbs, P.; Ward, R.L.; et al. Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Discov.* **2014**, *4*, 175–185.
3. Bielski, C.M.; Zehir, A.; Penson, A.V.; Donoghue, M.T.A.; Chatila, W.; Armenia, J.; Chang, M.T.; Schram, A.M.; Jonsson, P.; Bandlamudi, C.; et al. Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat. Genet.* **2018**, *50*, 1189–1195.

4. Taylor, A.M.; Shih, J.; Ha, G.; Gao, G.F.; Zhang, X.; Berger, A.C.; Schumacher, S.E.; Wang, C.; Hu, H.; Liu, J.; et al. Genomic and functional approaches to understanding cancer aneuploidy. *Cancer Cell* **2018**, *33*, 676–689.e3.
5. WHO Classification of Tumours Editorial Board. *Soft Tissue and Bone Tumours*, 5th ed.; IARC Press: Lyon, France, 2020.
6. Cancer Genome Atlas Research Network. Comprehensive and integrated genomic characterization of adult soft tissue sarcomas. *Cell* **2017**, *171*, 950–965.e28.
7. Soupir, A.; Ospina, O.E.; Hampton, O.; Churchman, M.; Radmacher, M.; Hedges, D.; McKean, D.; Agius, P.; Zeeshan, S.; Seligson, N.D.; et al. Genomic, transcriptomic, and immunogenomic landscape of over 1300 sarcomas of diverse histology subtypes. *Nat. Commun.* **2025**, *16*, 4206.
8. Cheng, D.T.; Mitchell, T.N.; Zehir, A.; Shah, R.H.; Benayed, R.; Syed, A.; Chandramohan, R.; Liu, Z.Y.; Won, H.H.; Scott, S.N.; et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J. Mol. Diagn.* **2015**, *17*, 251–264.
9. Nacev, B.A.; Sanchez-Vega, F.; Smith, S.A.; Antonescu, C.R.; Rosenbaum, E.; Shi, H.; Tang, C.; Socci, N.D.; Rana, S.; Gularte-Mérida, R.; et al. Clinical sequencing of soft tissue and bone sarcomas delineates diverse genomic landscapes and potential therapeutic targets. *Nat. Commun.* **2022**, *13*, 3405.
10. Schneider, D.; Gluski, J.; Brown, E.D.L.; Mishra, A.; Sciubba, D.M.; Lo, S.L. MOSAIC: A multi-feature genomic instability score independently predicts sarcoma survival in 2138 patients. *Am. J. Clin. Oncol.* **2025**, *48*, 465–469.
11. Chang, H.Y.; Dermawan, J.K.; Tap, W.; Singer, S.; Chi, P.; Antonescu, C.R. Novel genomic risk stratification model for primary high-grade malignant peripheral nerve sheath tumor (MPNST). *J. Pathol.* **2026**, *269*, 248–259.
12. Royston, P.; Altman, D.G.; Sauerbrei, W. Dichotomizing continuous predictors in multiple regression: A bad idea. *Stat. Med.* **2006**, *25*, 127–141.
13. Sauerbrei, W.; Taube, S.E.; McShane, L.M.; Cavenagh, M.M.; Altman, D.G. Reporting recommendations for tumor marker prognostic studies (REMARK): An abridged explanation and elaboration. *J. Natl. Cancer Inst.* **2018**, *110*, 803–811.
14. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2012**, *2*, 401–404. Correction in *Cancer Discov.* **2012**, *2*, 960.
15. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.E.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **2013**, *6*, pl1.
16. Shen, R.; Seshan, V.E. FACETS: Allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* **2016**, *44*, e131.
17. Carter, S.L.; Cibulskis, K.; Helman, E.; McKenna, A.; Shen, H.; Zack, T.; Laird, P.W.; Onofrio, R.C.; Winckler, W.; Weir, B.; et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* **2012**, *30*, 413–421.
18. Harrell, F.E., Jr.; Lee, K.L.; Mark, D.B. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* **1996**, *15*, 361–387.
19. DerSimonian, R.; Laird, N. Meta-analysis in clinical trials. *Control Clin. Trials* **1986**, *7*, 177–188.
20. Higgins, J.P.T.; Thompson, S.G.; Deeks, J.J.; Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ.* **2003**, *327*, 557–560.
21. Davidson-Pilon, C. Lifelines: Survival analysis in Python. *J. Open Source Softw.* **2019**, *4*, 1317.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.