

1 **Multicenter validation of an assay to predict anti-PD-1 disease control**
2 **in patients with recurrent or metastatic Head and Neck Squamous Cell**
3 **Carcinoma: The PREDAPT Study**

4 Kevin C. Flanagan¹, Jon Earls¹, Jeffrey Hiken¹, Rachel L. Wellinghoff¹, Michelle M.
5 Ponder¹, Howard L. Mcleod², William H. Westra³, Vera Vavinskaya⁴, Leisa Sutton⁵, Ida
6 Deichaite⁶, Orlan K. Macdonald⁷, Karim Welaya⁸, James L. Wade III⁹, Georges Azzi¹⁰,
7 Andrew W. Pippas¹¹, Jennifer Slim¹², Bruce Bank¹³, Xingwei Sui¹⁴, Steven E.
8 Kossman¹⁵, Todd D. Shenkenberg¹⁶, Warren L. Alexander¹⁷, Katharine A Price¹⁸,
9 Jessica Ley¹⁹, David N. Messina¹, Jarret I. Glasscock¹, A. Dimitrios Colevas²⁰, Ezra
10 E.W. Cohen⁴, Douglas R. Adkins¹⁹, Eric J. Duncavage²¹

11 ¹ Cofactor Genomics, Inc., St. Louis, MO, USA.

12 ² Intermountain Healthcare, St. George, UT, USA.

13 ³ Icahn School of Medicine at Mount Sinai, New York, New York, USA.

14 ⁴ Division of Hematology and Oncology, UCSD Moores Cancer Center, La Jolla, CA,
15 USA.

16 ⁵ UCSD Moores Cancer Center, La Jolla, CA, USA.

17 ⁶ Radiation Medicine & Applied Sciences, UCSD Moores Cancer Center, La Jolla, CA,
18 USA.

19 ⁷ Cancer Care Northwest, Spokane Valley, WA, USA.

20 ⁸ CoxHealth Medical Oncology, Springfield, MO, USA.

21 ⁹ Decatur Memorial Hospital, Decatur, IL, USA.

22 ¹⁰ Holy Cross Hospital Medical Group, Fort Lauderdale, FL, USA.

1 ¹¹ John B Amos Cancer Center, Columbus Regional Research Institute, Centricity
2 Research, Columbus, GA, USA.

3 ¹² Multicare Institute for Research and Innovation, Tacoma, WA, USA.

4 ¹³ Northwest Oncology and Hematology, Elk Grove Village, IL, USA.

5 ¹⁴ Providence Regional Cancer System, Lacey, WA, USA.

6 ¹⁵ Sharp Clinical Oncology Research, San Diego, CA, USA.

7 ¹⁶ Valley Cancer Associates, Harlingen, TX, USA.

8 ¹⁷ William Beaumont Army Medical Center and The Geneva Foundation, Fort Bliss, TX,
9 USA.

10 ¹⁸ Mayo Clinic, Rochester, MN, USA.

11 ¹⁹ Division of Oncology, Department of Medicine, Washington University School of
12 Medicine, St. Louis, MO, USA.

13 ²⁰ Stanford University, Stanford, CA USA.

14 ²¹ Department of Pathology and Immunology, Washington University School of
15 Medicine, St. Louis, MO, USA.

16

17 **ABSTRACT**

18 **Background:** Despite advances in cancer care and detection, more than 65% of
19 patients with squamous cell cancer of the head and neck (HNSCC) will develop
20 recurrent and/or metastatic disease. The prognosis for these patients is poor with a 5
21 year overall survival of 39%. Recent treatment advances in immunotherapy, including
22 immune checkpoint inhibitors like pembrolizumab and nivolumab, have resulted in
23 clinical benefit in a subset of patients. There is a critical clinical need to identify patients

1 who benefit from these anti-PD-1 immune checkpoint inhibitors. **Methods:** Here we
2 report findings from a multi-center observational study, PREDAPT (ClinicalTrials.gov:
3 NCT04510129), conducted across 17 US healthcare systems. PREDAPT aimed to
4 validate OncoPrism-HNSCC, a clinical biomarker assay predictive of disease control in
5 recurrent or metastatic HNSCC patients treated with anti-PD-1 immune checkpoint
6 inhibitors as a single agent (monotherapy) and in combination with chemotherapy
7 (chemo-immunotherapy). The test used RNA-sequencing data and machine learning
8 models to score each patient and place them into groups of Low, Medium, or High.
9 **Results:** The OncoPrism-HNSCC prediction significantly correlated with disease control
10 in both the monotherapy cohort (n=62, p=0.004) and the chemo-immunotherapy cohort
11 (n=50, p=0.01). OncoPrism-HNSCC also significantly predicted progression-free
12 survival in both cohorts (p=0.015 and p=0.037, respectively). OncoPrism-HNSCC had
13 more than threefold higher specificity than PD-L1 combined positive score and nearly
14 fourfold higher sensitivity than tumor mutational burden for predicting disease control.
15 **Conclusions:** Here we demonstrate the clinical validity of the OncoPrism-HNSCC
16 assay in identifying patients with disease control in response to anti-PD-1 immune
17 checkpoint inhibitors.

18

19 WHAT IS ALREADY KNOWN ON THIS TOPIC

20 Anti-PD-1 immune checkpoint inhibitors such as pembrolizumab benefit a subset of
21 patients with recurrent or metastatic head and neck squamous cell carcinoma (RM-
22 HNSCC), but current biomarkers are inadequate at identifying these patients.

1 WHAT THIS STUDY ADDS

2 This study describes the validation of a new RNA-based test that predicts disease
3 control and progression-free survival in response to anti-PD-1 therapy with high
4 sensitivity and specificity. The test was validated using two independent cohorts of
5 patients from 17 community and academic sites.

6 HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

7 The test had significantly higher sensitivity than TMB and significantly higher specificity
8 than PD-L1, enabling clinicians to make more informed decisions when prioritizing
9 treatment. Use of the test has the potential to avoid unnecessary chemotherapy and/or
10 anti-PD-1 treatment and improve patient outcomes.

11 **BACKGROUND**

12 Head and neck squamous cell carcinomas (HNSCC) represent a significant healthcare
13 burden. Worldwide, HNSCC is the seventh most common cancer with 870,000 new
14 cases and 440,000 deaths[1]. More than 65% of these HNSCC patients are ultimately
15 diagnosed with recurrent or metastatic disease[2,3]. Patients with recurrent or
16 metastatic head and neck squamous cell carcinoma (RM-HNSCC) have a poor
17 prognosis with median overall survival of just 10.7-13.0 months[4]. The introduction of
18 anti-PD-1 immune checkpoint inhibitors (ICIs) such as pembrolizumab and nivolumab
19 has improved outcomes for a subset of patients, but ICIs are associated with serious
20 adverse reactions and a high financial burden to the health system[5–9]. In practice,
21 many patients receive ICI in combination with platinum or other chemotherapies and

1 choosing between ICI monotherapy and ICI in combination with chemotherapy (chemo-
2 immunotherapy) is an important treatment decision.

3 Currently, anti-PD-1 treatments are recommended for patients with PD-L1
4 immunohistochemistry (IHC) combined positive score (CPS) of at least one. However,
5 while PD-L1 IHC testing is recommended for all patients, PD-L1 status does not
6 correlate well with clinical benefit, as measured by Disease Control Rate (DCR) defined
7 by the fraction of patients without disease progression as best tumor response post-
8 treatment. In KEYNOTE-048, for example, patients with PD-L1 CPS \geq 1 had only a 3%
9 higher DCR from immunotherapy as a single agent and a 1% lower DCR from chemo-
10 immunotherapy when compared to the same treatment arm regardless of PD-L1
11 CPS[4,10]. While Tumor Mutational Burden (TMB) is sometimes used to aid treatment
12 decisions, its clinical utility in HNSCC is less clear[11–13]. ICI undoubtedly provides
13 clinical benefit to a subset of patients, but current methods for identifying exactly which
14 patients benefit are insufficient.

15 There is an unmet clinical need for more robust methods of predicting disease control in
16 response to PD-1 inhibitors. Previously we described the development of an RNA-
17 sequencing-based classifier to predict disease control with increased sensitivity and
18 specificity compared to PD-L1 CPS in patients with RM-HNSCC treated with anti-PD-1
19 monotherapy[14]. In that study, we classified patients as progressors or non-
20 progressors based on the median predicted probability of disease control as determined
21 by the test. Building on that work, we refined the test and its thresholds for disease
22 control prediction to create three groups correlated with predicted likelihood of disease
23 control: Low, Medium, and High. Here we report the validation and test performance in

1 two independent cohorts of patients, one treated with anti-PD-1 alone (monotherapy)
2 and one with anti-PD-1 in combination with platinum-based chemotherapy (chemo-
3 immunotherapy). The resulting laboratory developed test (LDT), OncoPrism-HNSCC,
4 classifies patients into three groups, Low (25% of patients with lowest likelihood of
5 disease control), Medium (25% of patients with indeterminate likelihood of disease
6 control), and High (50% of patients with high likelihood of disease control). The test
7 predicts disease control in response to anti-PD-1 treatments with both high sensitivity
8 and specificity.

9 **METHODS**

10 **Study Design and Participants**

11 Patients were recruited from the following academic and community study sites across
12 the United States, with the aim of a representative sample of the affected population:
13 Washington University in St. Louis (St. Louis, MO), University of California San Diego
14 (San Diego, CA), Intermountain Healthcare (Salt Lake City, UT), Gundersen Medical
15 Foundation (La Crosse, WI), Cancer Care Northwest (Spokane, WA), Cox Medical
16 Centers (Springfield, MO), Decatur Memorial Hospital (Decatur, IL), Holy Cross Hospital
17 (Fort Lauderdale, FL), John B Amos Cancer Center (Columbus, GA), MultiCare Institute
18 for Research and Innovation (Tacoma, WA), Northwest Oncology and Hematology
19 (Hoffman Estates, IL), Ochsner Lafayette General Medical Center (Lafayette, LA),
20 Providence Regional Cancer System (Lacey, WA), Sharp Clinical Oncology Research
21 (San Diego, CA), Stanford University (Stanford, CA), William Beaumont Army Medical
22 Center (Fort Bliss, TX), Baylor College of Medicine (Houston, TX), Brooke Army Medical

1 Center (Fort Sam Houston, TX), Dayton Physicians Network (Dayton, OH), Mayo Clinic
2 (Rochester, MN), Revive Research Institute (Sterling Heights, MI), and Valley Cancer
3 Associates (Harlingen, TX).

4 Patients were enrolled from 2019-2023 in a combined prospective and retrospective,
5 observational study following the inclusion and exclusion criteria outlined below. No
6 patient level study data were reported to patients or physicians and the patients and
7 public were not involved in the study design. Eligible patients had recurrent or
8 metastatic histologically or cytologically confirmed HNSCC and were treated with
9 curative intent with anti-PD-1 either as a single agent (monotherapy) or in combination
10 with chemotherapy (chemo-immunotherapy) for recurrent or metastatic disease. Tissue
11 specimens analyzed in the study were collected from pre-treatment tumor samples
12 originally processed as formalin-fixed and paraffin embedded (FFPE) specimens using
13 standard histologic protocols. De-identified, pre-treatment FFPE tumor biopsy
14 specimens were provided to Cofactor Genomics for OncoPrism-HNSCC and PD-L1 IHC
15 analysis. Following treatment, each patient's tumor response to immunotherapy was
16 evaluated using RECIST, PERCIST, or other clinical criteria as appropriate in standard
17 of care to determine disease control. Patients with insufficient tissue for analysis (<10%
18 tumor cells as determined by a study pathologist [EJD]) and samples with greater than
19 22.4 months between biopsy and treatment were excluded from the study[14]. Primary
20 or metastatic tumor specimens were accepted, but metastatic tumors from liver or bone
21 were not included due to confounding tissue RNA expression and the difficulty of
22 recovering and processing decalcified FFPE RNA. Length of follow-up ranged from 34
23 days to 64 months. The study protocol, "A Multicenter Cancer Biospecimen Collection

1 Study” is registered as “NCT04510129—PREDicting immunotherapy efficacy from
2 Analysis of Pre-treatment Tumor biopsies (PREDAPT)” on clinicaltrials.gov. The study
3 protocol was approved by institutional review boards at either the study (Advarra, Inc.
4 [Columbia, MD] or WCG IRB [Puyallup, Washington]) or site level, as appropriate. All
5 patients provided signed, informed consent to participate, or consent was waived for
6 deceased patients according to study protocol. Independent data monitoring was
7 conducted by the study clinical research organization Curebase, Inc. (San Francisco,
8 CA).

9 **RNA Extraction**

10 RNA was extracted using RNAsform (Biotium, Fremont, CA) according to the
11 manufacturer’s instructions. RNA quantity was assessed by the High Sensitivity RNA
12 Qubit assay (Thermo Fisher Scientific, Waltham, MA). A predefined yield of 40 ng FFPE
13 RNA was used as the minimum QC threshold. Quality of the RNA was assessed using
14 a bioanalyzer (Agilent Technologies, Santa Clara, CA), and a DV200 of greater than
15 24% was required for all samples.

16 **Library Preparation and Sequencing**

17 Libraries were prepared using the QuantSeq 3’ mRNA-Seq Library Prep Kit FWD for
18 Illumina (Lexogen, Inc., Greenland, NH), following the manufacturer’s instructions.
19 Library RNA input was 40 ng for all samples. UMI Second Strand Synthesis Module for
20 QuantSeq FWD (Lexogen, Inc., Greenland, NH) replaced Second Strand Synthesis Mix
21 1 in the workflow. All samples were processed with two OncoPrism-HNSCC positive
22 controls and a No Template Control. The positive (high or medium scoring) controls

1 were RNA extracted from RM-HNSCC samples as described above. Final libraries were
2 sequenced to a minimum depth of 10 million single-end 75 base pair reads on a
3 NextSeq500 (Illumina, San Diego, CA), following the manufacturer's protocols.

4 **Immunohistochemistry**

5 PD-L1 staining was performed by Mosaic Labs (Lake Forest, CA) using the 22C3
6 pharmDx antibody (Agilent Technologies, Inc., Santa Clara, CA) or NeoGenomics
7 Laboratories (Fort Myers, FL) using the PD-L1 22C3 FDA (KEYTRUDA®) for HNSCC
8 Head and Neck stain. CPS assessment was performed by WHW or by NeoGenomics.
9 H&E staining was performed by NeoGenomics as part of the PD-L1 22C3 test or at
10 Cofactor Genomics using xylene substitute Slide Brite (Newcomer Supply, Middleton,
11 WI), as detailed by manufacturers.

12 **Processing of RNA sequencing data**

13 FASTQ files were preprocessed with trim_galore/cutadapt version 0.4.1 to remove
14 adapter sequences as well as reads with PHRED quality scores less than 20 and reads
15 that were less than 20 basepairs. The trimmed reads were aligned to the human
16 genome GRCh38 with STAR version 2.5.2a using the two-pass method as previously
17 described[15]. Read counts were generated using htseq-count version 0.9.1 and
18 annotation from Gencode version 22[15]. The data was normalized as counts per million
19 (CPM) and log₂ transformed using unique reads aligning to protein coding regions.
20 Samples were required to have a minimum of 30% exonic alignment and 800,000
21 unique deduplicated counts to be included in the study.

1 **Tumor Mutational Burden**

2 Tumor Mutational Burden (TMB) was measured using the GatewaySeq (Washington
3 University, St. Louis, MO) targeted DNA assay which was run in a CLIA-accredited
4 clinical laboratory. GatewaySeq calculates TMB using the Illumina (San Diego, CA)
5 Dragen TMB caller in tumor-only mode and non-synonymous TMB output. We used the
6 clinically validated GatewaySeq definition of TMB High (20 or more mutations per
7 megabase) to categorize patients as TMB High or TMB Low. Fifty to 250 ng DNA was
8 used as input. DNA was extracted using DNASTorm (Biotium, Fremont, CA) according to
9 the manufacturer's instructions.

10 **Model Training**

11 Data from 1205 total samples were used to select features, refine the protocol, train the
12 model, and validate the model (Figure 1). Data from 790 patient samples were used to
13 identify 149 candidate features related to immune response with detectable expression
14 across two publicly available datasets (Table S2)[16,17]. An additional 415 patient
15 samples were collected for the PREDAPT trial. Samples that were excluded or failed
16 QC requirements are detailed in Table S1. The remaining PREDAPT samples (n=211)
17 were divided into a training cohort and two validation cohorts based on treatment and
18 time enrollment was completed. The training cohort consisted of 99 samples from
19 patients receiving anti-PD-1 monotherapy at eleven PREDAPT healthcare systems[14].
20 A supervised machine learning, logistic regression model was built using this training
21 dataset. Patients with complete response, partial response or stable disease were

1 treated as the positive class. Samples from 34 patients ultimately assigned to validation
2 Cohort 1 were used as a preliminary evaluation of the training model performance.

3 **OncoPrism Scores and Prediction**

4 The OncoPrism-HNSCC biomarker generates an OncoPrism Score from 0 to 100 that
5 correlates with predicted disease control in patients with RM-HNSCC treated with anti-
6 PD-1 monotherapy. Higher OncoPrism Scores represent higher confidence by the
7 model that the patient will have disease control. The thresholds for the OncoPrism
8 Groups were defined from the training data. Considering n unique patient samples,
9 patients are chosen n times with replacement and used for training a model. Using this
10 trained model, a score ("out-of-bag score") is generated for the remainder of
11 patients[14,18]. This process was repeated 1000 times, and the out-of-bag score of
12 each patient was averaged to generate a mean training OncoPrism Score. The
13 threshold between the Low Group (OncoPrism Scores 0-37) and the Medium Group
14 (OncoPrism Scores 38-52) is defined as the value of the 25th percentile mean score.
15 The threshold between the Medium Group and the High Group (OncoPrism Scores 52-
16 100) is defined as the value of the 50th percentile mean score. These training cohort
17 mean score thresholds are used for all subsequent validation and analysis to define the
18 OncoPrism Groups.

19 **Validation of Performance**

20 Clinical validation of the OncoPrism-HNSCC assay was performed using a separate
21 cohort of 112 unique patient samples divided into two independent cohorts (Cohort 1
22 and Cohort 2). Samples were processed in the Cofactor Genomics CAP-accredited,

1 CLIA-certified laboratory using strict quality controls. The primary validation metric was
2 DCR in each OncoPrism Group. DCR was calculated by dividing the sum of patients
3 with RECIST 1.1-defined categories of stable disease, partial response, and complete
4 response as best response by the total number of patients in each group. DCR was
5 used because of similar PFS and clinical benefit previously observed among patients
6 with best response of stable disease and partial response[14]. To measure the test's
7 ability to enrich for disease control in response to anti-PD-1 monotherapy, 62 FFPE
8 tumor samples from 62 monotherapy treated patients from 15 clinical sites were
9 processed through the OncoPrism-HNSCC workflow (Cohort 1). As an additional
10 independent validation, 50 FFPE tumor samples from 50 chemo-immunotherapy treated
11 patients at 11 clinical sites were processed through the OncoPrism-HNSCC workflow
12 (Cohort 2). Patient specimens came from 17 unique clinical sites in total. OncoPrism
13 Scores were generated for each sample, and were assigned to the Low, Medium, or
14 High OncoPrism Groups based on these scores. Operators were blinded to the RECIST
15 label when processing samples and generating OncoPrism Scores. The RECIST labels
16 for each patient were determined independently from the OncoPrism Group and were
17 then used to determine the DCR for each group in the validation set.

18 For the purposes of test validation and treatment recommendations, Low Group patients
19 are classified as predicted progressors, Medium Group patients are considered an
20 indeterminate result, and High Group patients are classified as predicted to have
21 disease control in response to ICI. The Medium Group is considered an indeterminate
22 result due to the variation seen in Medium Group DCR and PFS across datasets (data
23 not shown and Fig 2B-C & E-F). When comparing performance with PD-L1 CPS, the

1 High Group patients are considered the predicted positive class (predicted disease
2 control), while the Low and Medium Group patients are treated as the predicted
3 negative class (no predicted disease control). Including the entire set of patients in
4 these calculations allows for a direct comparison to PD-L1 CPS even though the
5 intended use of OncoPrism-HNSCC is to consider a Medium Group result as
6 indeterminate.

7 **Statistics**

8 The primary endpoint of this study was disease control rate. A two-sided Cochran-
9 Armitage test for trends was used to test the significance of the trend of increasing
10 proportions for the DCRs of OncoPrism Groups. Power analysis was performed using
11 the training cohort out-of-bag area under the ROC curve, seeking a two-sided Type 1
12 error (alpha) of 0.05 and a Type 2 error (beta) of 0.80. A minimum sample size of 36
13 was calculated to power the primary endpoint (Cochran-Armitage Test for trend in
14 proportions for DCR). Expecting that training cohort out-of-bag performance may
15 overestimate independent cohort performance, we sought a minimum of 50 samples for
16 each cohort. PFS was defined as the time from start of ICI treatment to progression or
17 death. Patients were censored if they had not progressed at last follow up. One
18 OncoPrism High Group patient treated with chemo-immunotherapy was excluded
19 because of an unknown date of progression. PFS figures and analysis were done using
20 the “survminer” and “survival” packages, and significance was determined using log
21 rank methods [19,20]. Differences in sensitivity and specificity were tested using
22 McNemar’s test. 95% Confidence intervals for model performance metrics were

1 calculated using a non-parametric bootstrap resampling method. In all cases a p-value
 2 of less than 0.05 was considered significant.

3 **RESULTS**

4 Patients meeting all inclusion criteria (n=211) were divided into a training cohort (n=99)
 5 and two validation cohorts (Cohort 1, n=62 and Cohort 2, n=50; Figure 1; Table S3).
 6 The training and validation cohorts have similar patient and disease characteristics,
 7 except that the training cohort and validation Cohort 1 were treated with monotherapy
 8 ICI while validation Cohort 2 was treated with ICI in combination with chemotherapy
 9 (chemo-immunotherapy; Table 1).

10 Table 1: Patient Data

Characteristic	Designation	Training Set	Validation Cohort 1	Validation Cohort 2
Patient number	n	99	62	50
Age (median)	Age	67	69	60
Gender	Male	82%	69%	82%
	Female	17%	29%	18%
	unknown	1%	2%	0%
Smoking Status	Current or Former	72%	74%	74%
	Never	19%	23%	24%
	unknown	9%	3%	2%
ECOG	0 to 1	62%	50%	34%
	2 or greater	13%	11%	8%
	unknown	25%	39%	58%
Primary tumor site	Oral Cavity	30%	35%	44%
	Oropharynx	34%	31%	26%
	Larynx	14%	13%	12%
	Hypopharynx	4%	8%	4%
	Other/Unknown	17%	13%	14%
	Stage I, II	19%	26%	26%

Staging at diagnosis	Stage III, IV	71%	71%	64%
	Unknown	10%	3%	10%
HPV status	p16+	28%	19%	16%
	p16-/not tested	72%	80%	84%

1 Validation samples were processed with OncoPrism-HNSCC, generating an OncoPrism
 2 Score and resulting OncoPrism Group for each patient sample. Analytical variation was
 3 low, with highly repeatable results for replicate samples (manuscript in preparation). The
 4 primary endpoint of this study was disease control, so performance was evaluated using
 5 the DCR for each OncoPrism Group, with an expected trend from lower DCR in the Low
 6 Group to higher DCR in the High Group.

7 **Validation Cohort 1: monotherapy patients**

8 Specimens from 62 patients treated with anti-PD-1 monotherapy (Cohort 1) were scored
 9 and categorized into the Low, Medium, or High Group based on their score and the pre-
 10 established thresholds (Figure 2A). The groups roughly mirrored the expected
 11 population distribution, with 27% of patients in the Low Group, 32% of patients in the
 12 Medium Group, and 40% of patients in the High Group (Table 2).

13 Table 2: Clinical validation cohorts Disease Control Rates (DCR) by OncoPrism Group

OncoPrism Group	% of Cohort Population			Disease Control Rate (DCR)		
	Cohort 1*	Cohort 2*	Combined Cohort 1+2	Cohort 1 DCR	Cohort 2 DCR	Combined Cohort 1 & 2 DCR
Low	27%	22%	25%	12%	36%	21%
Medium	32%	32%	32%	55%	44%	50%
High	40%	46%	43%	60%	78%	69%
All	100%	100%	100%	45%	58%	51%

*Notes: % of Cohort Population = percent of subjects within the specified Group

1 In Cohort 1 the DCR increases from OncoPrism Low to Medium to High Groups (Table
2 2; Figure 2B; Cochran-Armitage trend $p=0.004$). In addition to higher DCR, patients in
3 the High Group also had significantly longer progression-free survival (PFS; log rank
4 test, $p=0.015$). Median PFS was 2.6 months for the Low Group, 4.2 months for the
5 Medium Group, and 9.8 months for the High Group (Figure 2C). Table 3 shows key
6 performance metrics when comparing the Low Group vs. the High Group, the two
7 actionable OncoPrism Groups. OncoPrism Medium patients were excluded from these
8 calculations as an indeterminate result (see Methods). OncoPrism-HNSCC predicted
9 disease control with high accuracy (0.71), high sensitivity (0.88), specificity (0.60),
10 positive predictive value (PPV [0.60]), and negative predictive value (NPV [0.88]).

11 Table 3: Performance metrics for validation cohorts

Cohort	Accuracy	Sensitivity	Specificity	PPV*	NPV*
Cohort 1: monotherapy (95% Confidence Interval) [Low vs High]	71% (58-86%)	88% (74-100%)	60% (39-83%)	60% (39-81%)	88% (75-100%)
Cohort 2: chemo-immunotherapy (95% Confidence Interval) [Low vs High]	74% (62-88%)	82% (65-96%)	58% (31-89%)	78% (62-95%)	64% (33-89%)

*Notes: PPV = Positive Predictive Power; NPV = Negative Predictive Power

12 Validation Cohort 2: chemo-immunotherapy patients

13 To test the ability of OncoPrism-HNSCC to predict disease control in these chemo-
14 immunotherapy-treated patients, performance of OncoPrism-HNSCC was evaluated in
15 specimens from 50 patients treated with anti-PD-1 in combination with chemotherapy
16 (Cohort 2). The groups roughly mirrored the expected population distribution, with 22%
17 of patients in the Low Group, 32% of patients in the Medium Group, and 46% of patients
18 in the High Group (Table 2; Figure 2D). As expected, the overall DCR for this cohort

1 was higher than the monotherapy cohort, likely due to the additional effect of
2 chemotherapy on outcome (58% vs. 45%; Table 2). As with Cohort 1, the DCR
3 increases from OncoPrism Low to Medium to High Groups (Table 2; Figure 2E;
4 Cochran-Armitage trend, $p=0.01$). This trend corresponded with significantly longer PFS
5 for patients in the High Group (log rank test, $p=0.037$). Median PFS was 3.0 months for
6 the Low Group, 3.4 months for the Medium Group, and 16.3 months for the High Group
7 in this cohort (Figure 2F). One patient from the High Group was excluded due to an
8 unknown date of progression ($n=49$). OncoPrism-HNSCC predicted disease control with
9 high accuracy (0.74), high sensitivity (0.82), specificity (0.58), PPV (0.78), and NPV
10 (0.64) when treating the OncoPrism High Group as the predicted positive class and the
11 OncoPrism Low Group as the predicted negative class (with the OncoPrism Medium
12 Group excluded as an indeterminate result; Table 3).

13 **OncoPrism-HNSCC is not predictive in non-ICI datasets**

14 Our data shows that the OncoPrism Group assignment is correlated with DCR in
15 patients treated with ICIs. To explore whether OncoPrism Group is predictive of disease
16 control in response to ICI or simply prognostic of outcome regardless of therapy, we
17 used the underlying OncoPrism-HNSCC model on four publicly available datasets of
18 HNSCC patients who were not treated with ICI [16,17,21,22]. The OncoPrism-HSNCC
19 biomarker was not significantly correlated with overall survival in any of the non-ICI
20 datasets (Table S4), suggesting that it is not an overall prognostic biomarker *per se* and
21 is consistent with the idea that OncoPrism-HSNCC is predictive of ICI disease control
22 specifically.

1 **OncoPrism-HNSCC outperforms the existing biomarkers PD-L1 CPS and TMB**

2 Currently, the biomarkers most frequently used to predict response to ICI in RM-
3 HNSCC patients are PD-L1 CPS and, less commonly, Tumor Mutational Burden (TMB).
4 Using our two validation cohorts, we compared the performance of OncoPrism-HNSCC
5 to PD-L1 CPS and TMB. First, we compared PD-L1 CPS to OncoPrism-HNSCC at all
6 possible thresholds for each biomarker using Receiver Operating Characteristic (ROC)
7 curves. For monotherapy-treated Cohort 1, the area under the curve (AUC) for
8 OncoPrism-HNSCC was 0.73, compared to 0.62 for PD-L1 CPS (Figure 3A). Likewise,
9 for the chemo-immunotherapy-treated Cohort 2, the OncoPrism-HNSCC AUC was 0.73
10 compared to 0.61 for PD-L1 CPS (Figure 3B).

11 Examining ROC curves is useful since each test has its own thresholds for dividing
12 groups. However, it is also important to compare performance using the commonly used
13 thresholds for each test. For OncoPrism-HNSCC, these thresholds are the divisions
14 between the Low, Medium, and High OncoPrism Groups. For PD-L1 CPS, we
15 categorized patients as PD-L1 Negative ($CPS < 1$) or PD-L1 Positive ($CPS \geq 1$), the
16 threshold recommended by ASCO guidelines[23]. In Cohort 1, the DCR for PD-L1
17 Negative patients was 25% compared to 48% for PD-L1 Positive patients. In Cohort 2,
18 PD-L1 Negative patients had a DCR of 50% compared to 59% for PD-L1 Positive. PD-
19 L1 status was not correlated with PFS in either cohort (Figure S1).

20 To compare sensitivity and specificity between OncoPrism-HNSCC and PD-L1 in the
21 same population, the OncoPrism High Group was designated as predicted disease
22 control (predicted positive class), while the OncoPrism Medium and Low Groups were

1 designated as predicted disease progression (predicted negative class). This strategy
2 differs from the metrics shown in Table 3, where the Medium Group was excluded as
3 indeterminate in order to match the intended use of the test, but it allows calculation of
4 metrics in the same patient populations for direct comparison. In the monotherapy
5 cohort (Figure 3C), OncoPrism-HNSCC had a sensitivity of 0.54 and a specificity of
6 0.71. The lower sensitivity compared to Table 3 is due to the inclusion of the Medium
7 Group. Using $CPS \geq 1$ to define predicted disease control, PD-L1 CPS had a sensitivity
8 of 0.93 in this cohort. However, the OncoPrism-HNSCC specificity of 0.71 is significantly
9 higher than the CPS specificity of 0.18 (McNemar's test, $p < 0.001$). Likewise, in the
10 chemo-immunotherapy cohort (Figure 3D), OncoPrism-HNSCC had a sensitivity of
11 0.62, compared to 0.90 for PD-L1 CPS (McNemar's test, $p < 0.05$). Again, the
12 OncoPrism-HNSCC specificity of 0.76 is significantly higher than the PD-L1 CPS
13 specificity of 0.14 (McNemar's test, $p < 0.001$; Figures 3C-D & Table S5). Reflecting the
14 relative sensitivities and specificities of each test, OncoPrism-HNSCC had more false
15 negatives while PD-L1 CPS had more false positives, although 71% of the OncoPrism-
16 HNSCC false negatives were in the Medium Group and would typically be treated as an
17 indeterminate result (Table S6). Interestingly, the proportions of patients in each CPS
18 category was very similar between Cohort 1 and Cohort 2, suggesting that the CPS
19 result was not influencing the treatment of the patients and skewing the datasets (Figure
20 3E).

21 Because PD-L1 CPS has low specificity, clinicians may have limited confidence that a
22 PD-L1 Positive patient will indeed benefit from ICI. To test whether OncoPrism-HNSCC
23 can predict PD-L1 Positive patients who have clinical benefit from ICI, we evaluated

1 PFS for each OncoPrism Group in PD-L1 Positive patients only (combined
2 monotherapy and chemo-immunotherapy cohorts). OncoPrism High patients had
3 significantly longer PFS than OncoPrism Medium or Low patients (log rank test,
4 $p < 0.001$; Figure S2).

5 TMB is less commonly used to guide treatment decisions in RM-HNSCC, but is
6 recommended in some tumor types and when PD-L1 CPS is not available[23]. To
7 compare the performance of OncoPrism-HNSCC to TMB, we evaluated TMB status for
8 samples from the monotherapy-treated cohort (Cohort 1). Specifically, all OncoPrism
9 High or Low samples with sufficient material were evaluated (32 samples in total). We
10 evaluated samples in the High and Low Groups as these two categories are the most
11 likely OncoPrism-HNSCC test results to influence a clinical decision. TMB of at least 20
12 mut/Mb was classified as TMB High, while less than 20 mut/Mb was considered TMB
13 Low (see Methods). Overall, OncoPrism-HNSCC had a sensitivity of 0.85 and a
14 specificity of 0.53 in this group, compared to a sensitivity of 0.23 and a specificity of 1
15 for TMB (Figure 3F). The sensitivity and specificity for CPS in this group is also shown
16 for reference. While TMB had significantly higher specificity than OncoPrism-HNSCC
17 (McNemar's test, $p = 0.008$), it only identified three patients with disease control (Table
18 S7). OncoPrism-HNSCC had significantly higher sensitivity than TMB (McNemar's test,
19 $p = 0.027$).

20 **DISCUSSION**

21 OncoPrism-HNSCC significantly predicts disease control and PFS in response to anti-
22 PD-1 (ICI) therapy in pre-treatment RM-HNSCC patients. Importantly, the test was

1 validated in two separate cohorts using patient samples from 17 clinical academic and
2 community sites from across the United States, which allowed us to account for test
3 performance across a variety of possible pre-analytic sample processing conditions.
4 The multi-dimensional biomarker underlying OncoPrism-HNSCC was built using the
5 careful evaluation of a previously published study of cell composition, cell-state, and
6 immune modulatory genes in the tumor microenvironment [14]. Both validation cohorts
7 (Cohort 1: monotherapy and Cohort 2: chemo-immunotherapy) had similar results, with
8 a significant correlation of OncoPrism Group classification with DCR and PFS, as well
9 as high accuracy, sensitivity, specificity, PPV, and NPV. The OncoPrism-HNSCC model
10 was not predictive in patients treated with non-ICI therapies, suggesting that the
11 biomarker is predictive rather than prognostic (Table S4). These results also suggest
12 that the predictive nature of the biomarker may be specific to the ICI component in
13 chemo-immunotherapy-treated Cohort 2.

14 There was no distinct difference observed in PD-L1 CPS status of those patients
15 prescribed monotherapy (Cohort 1) vs. chemo-immunotherapy (Cohort 2) (Figure 3E),
16 suggesting that PD-L1 score was not driving treatment decision between monotherapy
17 and chemo-immunotherapy in this study. Limitations of PD-L1 for guiding treatment
18 have been previously published[24,25].

19 The intended use of OncoPrism-HNSCC is to aid clinicians in choosing whether to treat
20 with anti-PD-1 as a single agent, anti-PD-1 in combination with chemotherapy, or
21 alternative treatment options. While patients treated with chemo-immunotherapy have
22 higher DCR than monotherapy-treated patients, chemo-immunotherapy is also
23 associated with higher toxicity[4,9,26]. OncoPrism-HNSCC improves upon other assays

1 to inform these treatment decisions, with the potential to limit unnecessary treatment-
2 associated toxicities and improve patient outcomes. Currently, PD-L1 CPS is the most
3 common biomarker used to guide such decisions. Unfortunately, PD-L1 CPS has high
4 sensitivity but low specificity for predicting disease control (Fig 3C-D and Table S5).
5 This low specificity means that many patients with high CPS do not clinically benefit
6 from ICI, and clinicians are reluctant to use the CPS to exclude patients from more
7 aggressive treatment options like chemo-immunotherapy. OncoPrism-HNSCC has
8 significantly higher specificity than PD-L1. In addition, OncoPrism Groups stratify
9 patients by PFS among all patients (Figure 2C&F) and in PD-L1 Positive patients alone
10 (Figure S2). Together, these results suggest that ICI therapy should be prioritized for
11 OncoPrism High patients (Figure 4). Given the low DCR and PFS in the OncoPrism Low
12 Group, clinicians should consider a non-ICI treatment and/or available clinical trials for
13 OncoPrism-HNSCC Low patients regardless of PD-L1 status. While PD-L1 CPS had
14 significantly higher sensitivity than OncoPrism-HNSCC when categorizing OncoPrism
15 Low and Medium patients as predicted progressors (Figure 3C-D), OncoPrism-HNSCC
16 had similar sensitivity to PD-L1 when comparing OncoPrism Low to High (see Methods;
17 Table 3 and Table S5). In contrast to PD-L1, TMB had high specificity but low sensitivity
18 (Figure 3F). As a result, ICI treatment should be prioritized for TMB High patients, but
19 OncoPrism-HNSCC results appear to have superior prediction over TMB Low results.
20 Figure S3 provides a summary of treatment recommendations based on test results.

21 This real-world observational validation study has several limitations. These limitations
22 include patients with long time between test biopsy and treatment, the inclusion of
23 patients who had intervening treatments between the biopsy and the ICI, and instances

1 of incomplete data within the patient’s clinical record. The inclusion criteria in this study
2 balanced recruiting real-world, well-controlled patient cohorts reflecting the intended
3 treatment scenario with maximizing the study size. Ongoing and future studies aim to
4 refine the model and study additional endpoints such as overall survival with additional
5 patient cohorts. In addition, future work will refine the thresholds for each OncoPrism
6 Group to maximize the number of patients with actionable results without sacrificing the
7 sensitivity or specificity of the test. The Medium Group is considered an indeterminate
8 result due to the variation seen in Medium Group DCR and PFS across the training
9 cohort and validation cohorts (data not shown and Fig 2B-C & E-F). As a result,
10 currently the Medium Group result does not provide clear clinical guidance to physicians
11 and patients. In the validation cohorts, 32% of patients fell in the Medium Group,
12 meaning 68% of patients would have received treatment-guiding information. Because
13 OncoPrism-HNSCC balances high sensitivity and high specificity for predicting disease
14 control, it is clinically useful and has the potential to aid treatment decisions of more
15 patients than existing tests.

16 Although there is no ICI predictive biomarker with perfect sensitivity and specificity,
17 OncoPrism-HNSCC addresses significant shortcomings of PD-L1 and TMB in the RM-
18 HNSCC population through a balance of sensitivity and specificity, enabling clinicians to
19 identify patients most likely to benefit from immunotherapy. OncoPrism-HNSCC exhibits
20 clinical validity across a diverse patient population and holds promise to guide treatment
21 decisions and improve patient outcomes.

22 **DECLARATIONS**

23 **Ethics approval and consent to participate**

1 The study protocol was approved through WCG or Advarra IRB and/or at the local site
2 level IRB. Independent data monitoring was conducted by the study clinical research
3 organization Curebase, Inc (San Francisco, CA). All patients provided signed, informed
4 consent to participate, or consent was waived for deceased patients according to the
5 approved study protocol.

6 **Consent for publication**

7 Not applicable

8 **Availability of data and material**

9 The data underlying this study, including anonymized patient-level OncoPrism-HNSCC,
10 PD-L1, and TMB measurements, are available for non-commercial use from the
11 corresponding author upon reasonable request. Materials and analytical code are not
12 available.

13 **Competing interests**

14 KCF, JE, JH, RLW, MMP, DNM, JIG, and EJD are employed, have stock interests,
15 and/or a financial relationship with Cofactor Genomics, Inc., maker of the OncoPrism-
16 HSNCC test.

17 **Funding**

18 This study was wholly funded by Cofactor Genomics, Inc.

19 **Authors' contributions**

20 Conception and design: KCF, JE, JH, RLW, MMP, DNM, JIG, EJD

21 Recruitment of patients and acquisition of samples and data: MMP, HM, VV, LS, ID,

22 OKM, KW, JLW, GA, AWP, JS, BB, XS, SEK, TDS, WLA, KAP, JL, DNM, JIG, ADC,

23 EEWC, DRA,

1 Execution of the research: KCF, JE, JH, RLW, MMP, WHW, EJD

2 Analysis and interpretation of data: KCF, JE, JH, EJD

3 Drafting of the manuscript: KCF, JE, JIG

4

5 **ACKNOWLEDGMENTS**

6 We would like to thank the patients who participated in this study. In addition, we thank
7 all members of the PREDAPT study consortium. In particular, we thank the following
8 people for their invaluable contributions: Courtney Alexander, Rachel Allen, Newsheen
9 Azeemuddin, Rajesh Bande, Brandon Barker, Savanna Biedermann, Darcie A. Cruz,
10 Karl D'Silva, Paula Datri, Deirdre L. Dillon, Julie A. Faust, Eileen M. Georgi, Paul
11 Gonzales, Carrie Grandier, Leah A. Guilford-Elenes, Ann Hargrove, Peter Jiang,
12 Paraic A. Kenny, Cheryl L. Krantz, Mukesh Kumar, Natalie K. Kyek, Steven Lai, Amber
13 N. Larimer, Jenna R. Larry, Robert C. Matulonis, Elyssa Navarro, Uyen Nguyen, Suresh
14 Nukala, Lucyna Olechny, Meley L. Pine, Breanna N. Richardson, Julie Roache, Destiny
15 Robinson, Steven Saccaro, Alaina B. Sandoz, Kelly Shrimpton, Julie Skarsvog, Peggy
16 Smith, Scott A. Sonnier, Galen A. Steinhoff, Jennifer Taylor, Alierykel Williams, Erin L.
17 Wilson, and Anna Young.

18

19 **KEYWORDS**

20 HNSCC; biomarker; immune checkpoint inhibitors; PD-L1; Tumor Mutational Burden;
21 pembrolizumab; nivolumab

22

23 **REFERENCES**

- 1 1 Sung H, Ferlay J, Siegel RL, *et al.* Global Cancer Statistics 2020: GLOBOCAN
2 Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.
3 *CA Cancer J Clin.* 2021;71:209–49.
- 4 2 Argiris A, Karamouzis M V, Raben D, *et al.* Head and neck cancer. *The Lancet.*
5 2008;371:1695–709.
- 6 3 Chow LQM. Head and neck cancer. *N Engl J Med.* 2020;382:60–72.
- 7 4 Burtneess B, Harrington KJ, Greil R, *et al.* Pembrolizumab alone or with
8 chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic
9 squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised,
10 open-label, phase 3 study. *The Lancet.* 2019;394:1915–28.
- 11 5 Larkins E, Blumenthal GM, Yuan W, *et al.* FDA Approval Summary:
12 Pembrolizumab for the Treatment of Recurrent or Metastatic Head and Neck
13 Squamous Cell Carcinoma with Disease Progression on or After Platinum-
14 Containing Chemotherapy. *Oncologist.* 2017;22:873–8.
- 15 6 Ferris RL, Blumenschein G, Fayette J, *et al.* Nivolumab for Recurrent Squamous-
16 Cell Carcinoma of the Head and Neck. *New England Journal of Medicine.*
17 2016;375:1856–67.
- 18 7 Muro K, Chung HC, Shankaran V, *et al.* Pembrolizumab for patients with PD-L1-
19 positive advanced gastric cancer (KEYNOTE-012): a multicentre, open-label,
20 phase 1b trial. *Lancet Oncol.* 2016;17:717–26.
- 21 8 Wang H, Mustafa A, Liu S, *et al.* Immune checkpoint inhibitor toxicity in head and
22 neck cancer: From identification to management. *Front Pharmacol.* 2019;10.
23 <https://doi.org/10.3389/fphar.2019.01254>
- 24 9 Kessler R, Pandruvada S. Immune-related adverse events following checkpoint
25 inhibitor treatment in head and neck cancers: A comprehensive review. *Oral*
26 *Oncology Reports.* 2023;6. <https://doi.org/10.1016/j.oor.2023.100036>
- 27 10 Harrington KJ, Burtneess B, Greil ; Richard, *et al.* Pembrolizumab With or Without
28 Chemotherapy in Recurrent or Metastatic Head and Neck Squamous Cell
29 Carcinoma: Updated Results of the Phase III KEYNOTE-048 Study. *J Clin Oncol.*
30 2022;41:790–802.
- 31 11 Aggarwal C, Ben-Shachar R, Gao Y, *et al.* Assessment of Tumor Mutational
32 Burden and Outcomes in Patients With Diverse Advanced Cancers Treated With
33 Immunotherapy. *JAMA Netw Open.* 2023;6:e2311181.
- 34 12 McGrail DJ, Pilié PG, Rashid NU, *et al.* High tumor mutation burden fails to
35 predict immune checkpoint blockade response across all cancer types. *Annals of*
36 *Oncology.* 2021;32:661–72.
- 37 13 Haddad RI, Seiwert TY, Chow LQM, *et al.* Influence of tumor mutational burden,
38 inflammatory gene expression profile, and PD-L1 expression on response to
39 pembrolizumab in head and neck squamous cell carcinoma. *J Immunother*
40 *Cancer.* 2022;10. doi: 10.1136/jitc-2021-003026
- 41 14 Flanagan KC, Earls J, Schillebeeckx I, *et al.* Multidimensional biomarker predicts
42 disease control in response to immunotherapy in recurrent or metastatic head and
43 neck squamous-cell carcinoma. *J Cancer Res Clin Oncol.* 2023;149:14125–36.
- 44 15 Schillebeeckx I, Armstrong JR, Forys JT, *et al.* Analytical Performance of an
45 Immunoprofiling Assay Based on RNA Models. *Journal of Molecular Diagnostics.*
46 2020;22:555–70.

- 1 16 Lawrence MS, Sougnez C, Lichtenstein L, *et al.* Comprehensive genomic
2 characterization of head and neck squamous cell carcinomas. *Nature*.
3 2015;517:576–82.
- 4 17 Wichmann G, Rosolowski M, Krohn K, *et al.* The role of HPV RNA transcription,
5 immune response-related gene expression and disruptive TP53 mutations in
6 diagnostic and prognostic profiling of head and neck cancer. *Int J Cancer*.
7 2015;137:2846–57.
- 8 18 B Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of*
9 *Statistics*. 1979;7:1–26.
- 10 19 Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model.
11 Published Online First: 2000. doi: 10.1007/978-1-4757-3294-8
- 12 20 Therneau TM. Survival Analysis [R package survival version 3.5-5]. Published
13 Online First: 12 March 2023.
- 14 21 Keck MK, Zuo Z, Khattri A, *et al.* Integrative analysis of head and neck cancer
15 identifies two biologically distinct HPV and three non-HPV subtypes. *Clinical*
16 *Cancer Research*. 2015;21:870–81.
- 17 22 Lohavanichbutr P, Méndez E, Holsinger FC, *et al.* A 13-gene signature prognostic
18 of HPV-negative OSCC: Discovery and external validation. *Clinical Cancer*
19 *Research*. 2013;19:1197–203.
- 20 23 Yilmaz E, Ismaila N, Bauman JE, *et al.* Immunotherapy and Biomarker Testing in
21 Recurrent and Metastatic Head and Neck Cancers: ASCO Guideline. *J Clin*
22 *Oncol*. 2022;41:1132–46.
- 23 24 Liu X, Guo C-Y, Tou F-F, *et al.* Association of PD-L1 expression status with the
24 efficacy of PD-1/PD-L1 inhibitors and overall survival in solid tumours: A
25 systematic review and meta-analysis. *Int J Cancer*. 2020;147:116–27.
- 26 25 Park JC, Krishnakumar HN, Saladi SV. Current and Future Biomarkers for
27 Immune Checkpoint Inhibitors in Head and Neck Squamous Cell Carcinoma.
28 *Current Oncology*. 2022;29:4185–98. <https://doi.org/10.3390/curroncol29060334>
- 29 26 Cohen EEW, Soulières D, Le Tourneau C, *et al.* Pembrolizumab versus
30 methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck
31 squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3
32 study. *The Lancet*. 2019;393:156–67.
- 33

34 **FIGURE LEGENDS**

35 **Figure 1.** Samples used to develop, train, and validate OncoPrism-HNSCC. Data from
36 a total of 1205 samples were used to select features, refine the protocol, train the
37 model, and validate the model. Data from 790 publicly available samples were used to
38 select features. 415 patient samples were collected as part of the PREDAPT trial. 116
39 samples were ineligible and were excluded from this study. 86 samples failed quality

1 control and were excluded. Two patients were withdrawn from the study. Of the
2 remaining patient samples, 161 were treated with monotherapy anti-PD-1 and 50 were
3 treated with chemo-immunotherapy anti-PD-1. Of the monotherapy samples, 99 were
4 used to train the model, and the remaining 62 monotherapy samples served as
5 validation Cohort 1. The 50 chemo-immunotherapy samples served as validation Cohort
6 2.

7
8 **Figure 2.** OncoPrism-HNSCC Score and Group are correlated with disease control in
9 independent monotherapy and chemo-immunotherapy validation cohorts. Samples are
10 ordered by their OncoPrism Score for the monotherapy-treated (**A**) and chemo-
11 immunotherapy (**D**) validation cohorts. Lower scores are more likely to be progressors
12 (grey) while higher scores are more likely to have disease control (orange). Based on
13 their OncoPrism Score and predetermined thresholds (dotted lines), each patient
14 sample is assigned to an OncoPrism Group (Low, Medium, or High). OncoPrism
15 Groups are significantly correlated with disease control rate (DCR) in the monotherapy
16 ($p=0.004$) (**B**) and chemo-immunotherapy ($p=0.01$) (**E**) validation cohorts. P-values for
17 the significance of the trend were calculated using Cochran-Armitage. OncoPrism
18 Groups are significantly correlated with Progression-Free Survival (PFS) in the
19 monotherapy ($p=0.015$) (**C**) and chemo-immunotherapy ($p=0.037$) (**F**) validation
20 cohorts. P-values for PFS were calculated using log rank methods.

21
22 **Figure 3.** OncoPrism-HNSCC outperforms existing assays PD-L1 CPS and TMB.
23 Receiver Operating Characteristic (ROC) curves are shown for the monotherapy (**A**)

1 and chemo-immunotherapy **(B)** cohorts. OncoPrism-HNSCC (orange) has a higher area
2 under the curve (AUC) than PD-L1 CPS (grey) in both cohorts. In monotherapy **(C)** and
3 chemo-immunotherapy **(D)** cohorts, OncoPrism-HNSCC (orange) has high sensitivity
4 and specificity, while PD-L1 CPS (grey) has high sensitivity but low specificity. The
5 distribution of CPS is similar in each validation cohort **(E)**. **(F)** Sensitivity and specificity
6 for OncoPrism-HNSCC (orange), PD-L1 CPS (grey), and TMB (blue) in 32 patients from
7 the OncoPrism High and Low Groups. Error bars represent 95% Confidence Intervals.

8
9 **Figure 4.** Immune checkpoint inhibitor decision tree based on test results. Patients with
10 recurrent or metastatic HNSCC tested with OncoPrism-HNSCC are categorized into the
11 OncoPrism Low, Medium, or High Group. Because OncoPrism-HNSCC has high
12 specificity relative to PD-L1 CPS and OncoPrism High patients have longer PFS
13 regardless of PD-L1 status, OncoPrism High patients should typically be treated with ICI
14 regardless of PD-L1 status. OncoPrism Low patients have low ICI DCR and are not
15 good candidates for ICI even if they are PD-L1 Positive. Patients in the OncoPrism
16 Medium Group do not have a definitive treatment path; all test results and treatment
17 options should be considered. Typically, ICI should be favored for OncoPrism Medium
18 patients who are PD-L1 Positive while non-ICI or clinical trial options should be
19 considered for PD-L1 Negative patients. TMB testing is not recommended for most
20 HNSCC patients. However, if TMB testing is performed, ICIs should be prioritized for
21 TMB High patients given the high observed specificity of TMB. Only 9% of patients in
22 our study were TMB High. Due to the low sensitivity of TMB, a TMB Low result should
23 not be strongly considered in treatment decisions.

1

2 **Figure S1.** PD-L1 does not predict progression-free survival (PFS). Patients were
3 divided into PD-L1 Positive (CPS \geq 1) or Negative (CPS $<$ 1) groups within each validation
4 cohort. PD-L1 status was not significantly correlated with PFS in the monotherapy
5 (p=0.71) **(A)** or chemo-immunotherapy (p=0.28) **(F)** validation cohorts. P-values for PFS
6 were calculated using log rank methods.

7

8 **Figure S2.** OncoPrism-HNSCC predicts progression-free survival (PFS) in PD-L1
9 Positive patients. All PD-L1 Positive patients from both Cohort 1(monotherapy) and
10 Cohort 2 (chemo-immunotherapy) were grouped according to their OncoPrism-HNSCC
11 result. OncoPrism High Group is significantly associated with longer PFS (p $<$ 0.001, log
12 rank methods).

13

14 **Figure S3.** Treatment prioritization for discordant test results. The recommended
15 treatment for concordant (grey) and discordant (white) test results is shown for
16 OncoPrism-HNSCC compared to PD-L1 CPS **(A)** or TMB **(B)**. The percent shown is the
17 percentage of the study population in each category. The most common discordant test
18 results are outlined for PD-L1 (red) and TMB (blue). OncoPrism Medium samples were
19 considered indeterminate and are not included. For comparison with PD-L1, n=75. For
20 comparison with TMB, n=32.

Figure 1

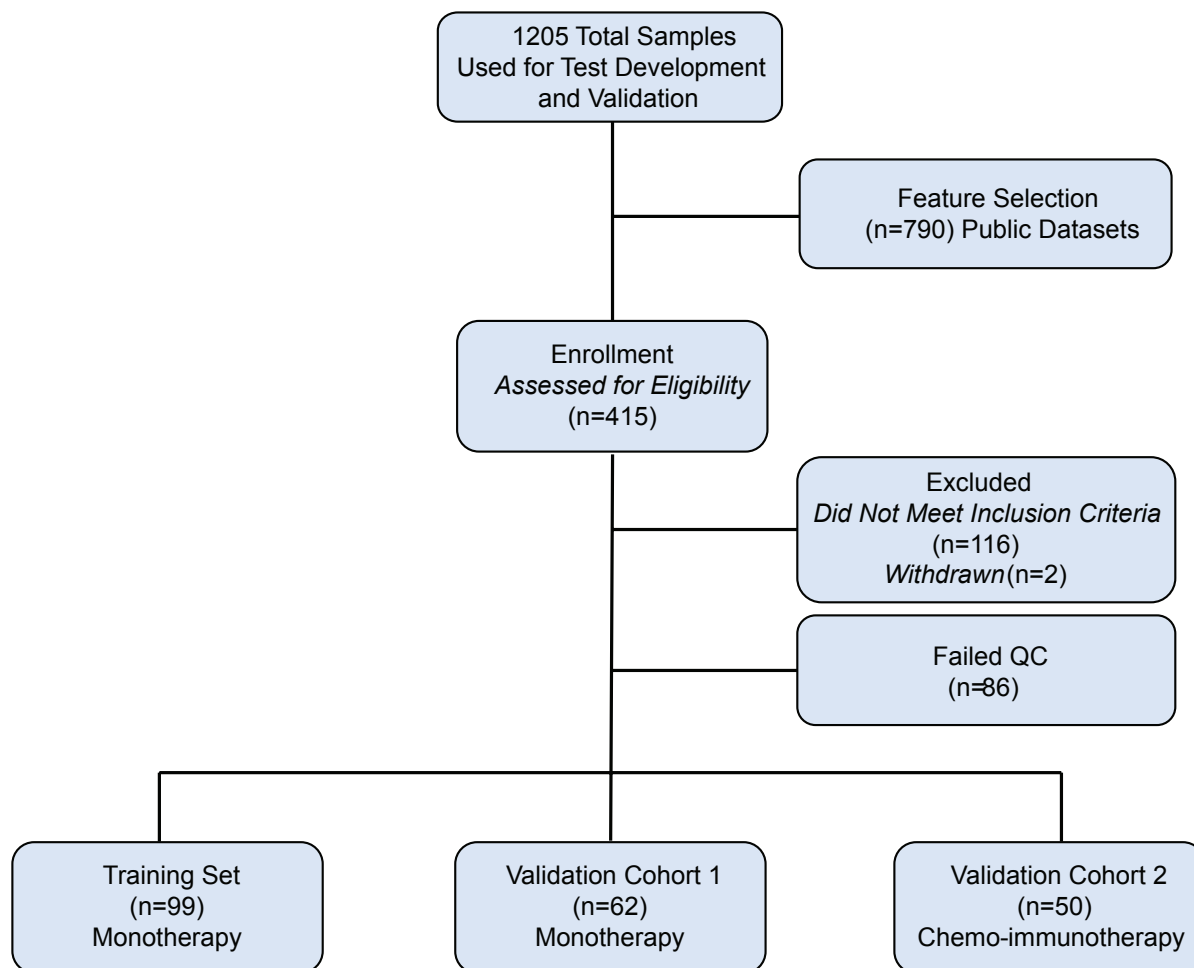


Figure 2

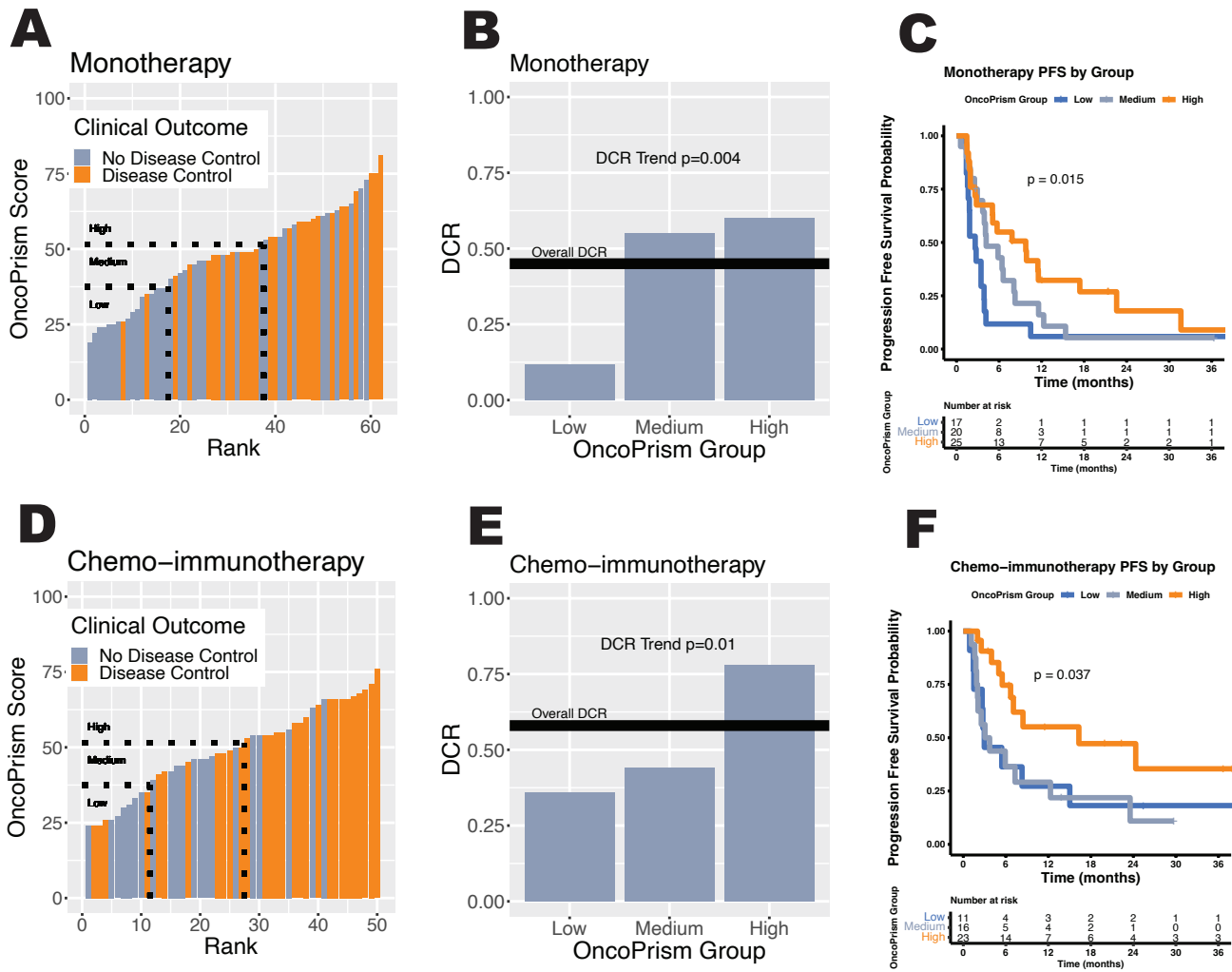


Figure 3

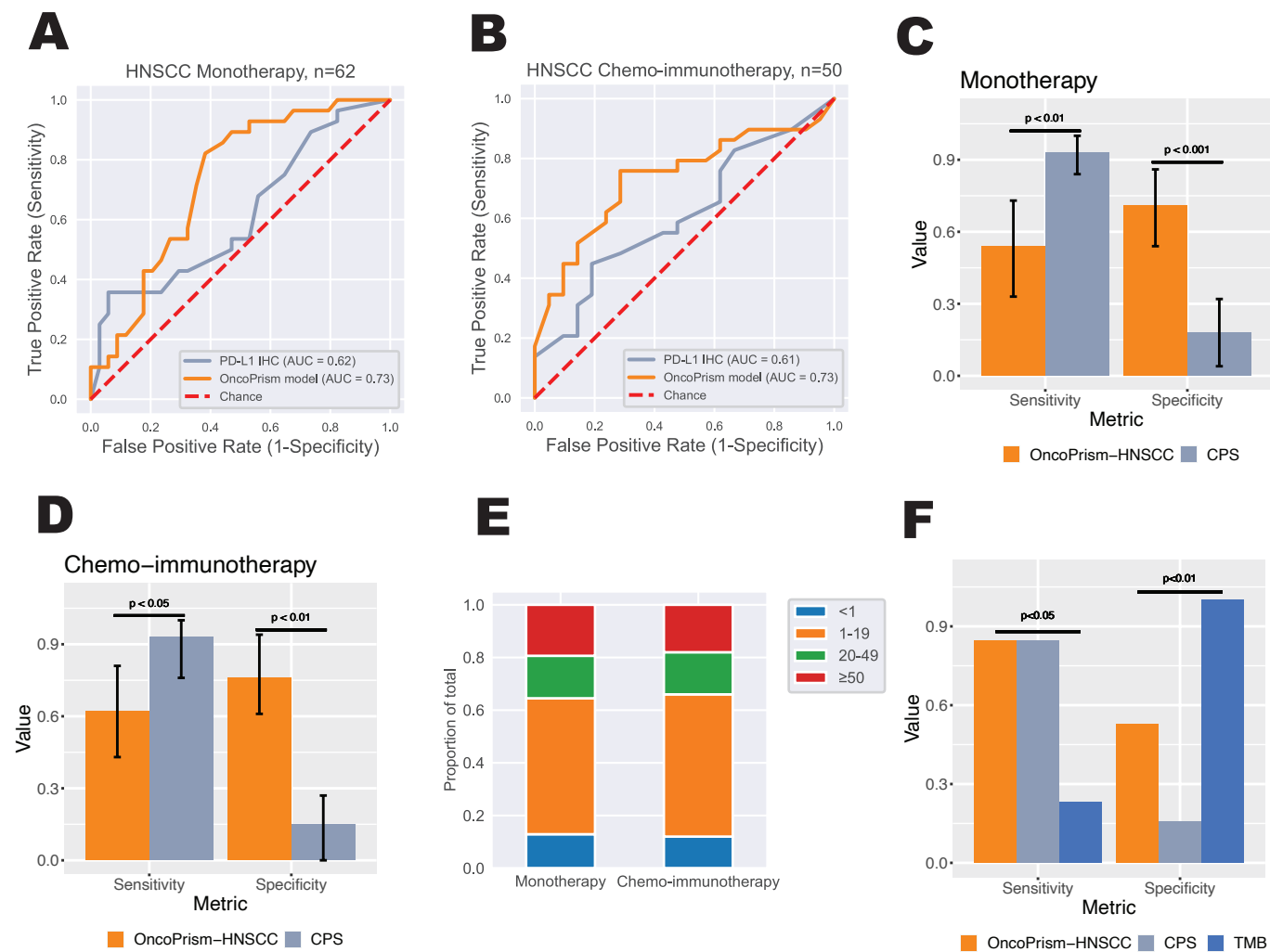


Figure 4

