It is made available under a CC-BY 4.0 International license .

# 1 Predicting Progression Free Survival after Systemic Therapy in Advanced

# 2 Head and Neck Cancer: Bayesian regression and Model development

Paul R Barber<sup>‡1,2</sup>, Fabian Flores-Borja<sup>‡\*3</sup>, Giovanna Alfano<sup>‡4</sup>, Kenrick Ng<sup>‡1</sup>, Gregory Weitsman<sup>4</sup>,
Luigi Dolcetti<sup>4</sup>, Rami Mustapha<sup>4</sup>, Felix Wong<sup>4</sup>, Jose M Vicencio<sup>1,4</sup>, Myria Galazi<sup>1</sup>, James W

5 Opzoomer<sup>5</sup>, James N Arnold<sup>5</sup>, Shahram Kordasti<sup>6</sup>, Jana Doyle<sup>7</sup>, Jon Greenberg<sup>7</sup>, Magnus T

6 Dillon<sup>8</sup>, Kevin J Harrington<sup>8</sup>, Martin D Forster<sup>1</sup>, Anthony C C Coolen<sup>9,10</sup>, and Tony Ng<sup>1,3,4†</sup>.

- 8 1. UCL Cancer Institute, Paul O'Gorman Building, University College London, London, UK
- 9 2. Comprehensive Cancer Centre, School of Cancer & Pharmaceutical Sciences, King's10 College London, London, UK.
- Breast Cancer Now Research Unit, School of Cancer & Pharmaceutical Sciences, King's
   College London, London, UK.
- Richard Dimbleby Laboratory of Cancer Research, School of Cancer & Pharmaceutical
   Sciences, King's College London, London, UK.
- 15 5. Tumor Immunology Group, School of Cancer & Pharmaceutical Sciences, King's College
  16 London, London, UK
- Systems Cancer Immunology, School of Cancer & Pharmaceutical Sciences, King's College
   London, London, UK
- 19 7. Daichii Sankyo Incorporated, New Jersey, USA
- 20 8. The Institute of Cancer Research, London, UK
- 21 9. Institute for Mathematical and Molecular Biomedicine, King's College London, UK
- 22 10. Saddle Point Science Ltd, London, UK
- 23
- 24 ‡ These authors contributed equally to this work.
- 25 \* Current address: Centre for Immunobiology and Regenerative Medicine, Barts & The
- 26 London School of Medicine and Dentistry, Queen Mary University of London, London, UK
- <sup>†</sup>To whom correspondence should be addressed:
- 28 Prof. Tony Ng

It is made available under a CC-BY 4.0 International license .

- 29 Tel no: +44 (0) 20 7848 8056 (KCL)
- 30 E-mail: tony.ng@kcl.ac.uk, <u>t.ng@ucl.ac.uk</u>
- 32 Word Count: 4833 words (excluding figure legends and references)
- **Abstract:** 248 words
- 34 Running Title: Predicting PFS in Head and Neck Cancer

It is made available under a CC-BY 4.0 International license .

# 49 List of abbreviations:

50	C1 or C2	Timepoint before cycle 1 or cycle 2 of treatment
51	C-index	Harrell's Concordance Index
52	СТ	Computed Tomography
53	CyTOF	Mass Cytometry
54	ddPCR	Digital Droplet Polymerase Chain Reaction
55	DN	Double-Negative
56	EGFR	Epidermal Growth Factor Receptor
57	FLIM	Fluorescence Lifetime Imaging Microscopy
58	FRET	Förster Resonance Energy Transfer
59	HPV	Human Papilloma Virus
60	HNSCC	Head and Neck Squamous Cell Cancer
61	HR	Hazard Ratio
62	KM	Kaplan-Meier
63	LFC	Log Fold Change
64	miRNA	Micro-RNA
65	PBMC	Peripheral Blood Mononuclear Cells
66	PCR	Polymerase Chain Reaction
67	PFS	Progression Free Survival
68	Tregs	Regulatory T cells

It is made available under a CC-BY 4.0 International license .

#### 69 ABSTRACT

#### 70 Background:

Advanced Head and Neck Squamous Cell Cancer (HNSCC) is associated with a poor
prognosis, and biomarkers that predict response to treatment are highly desirable. The primary
aim was to predict Progression Free Survival (PFS) with a multivariate risk prediction model.

74 Methods:

Blood samples from 56 HNSCC patients were prospectively obtained within a Phase 2 clinical trial (NCT02633800), before and after the first treatment cycle of platinum-based chemotherapy, to identify biological covariates predictive of outcome. A total of 42 baseline covariates were derived pre-treatment, which were combined with 29 covariates after one cycle of treatment. These covariates were ranked and selected by Bayesian multivariate regression to form risk scores to predict PFS, producing "baseline" and "combined" risk prediction models respectively.

81 Results:

82 The baseline model comprised of CD33+CD14+ monocytes, Double Negative B cells and age, 83 in a weighted risk signature which predicted PFS with a concordance index (C-index) of 0.661. 84 The combined model composed of baseline CD33+CD14+ monocytes, baseline Tregs, after-85 treatment changes in CD8 effector memory T cells, CD8 Central memory T cells and CD3 T 86 Cells, along with the hypopharyngeal primary tumor subsite. This weighted risk signature 87 exhibited an improved C-index of 0.757. There was concordance between levels of CD33+CD14+ myeloid cells in tumor tissue, as demonstrated by imaging mass cytometry, and 88 peripheral blood in the same patients. This monocyte subpopulation also had univariate 89 90 predictive value (log-rank p value = 0.03) but the C-index was inferior to the combined 91 signature.

92 Conclusions:

93 This immune-based combined multimodality signature, obtained through longitudinal peripheral blood monitoring, presents a novel means of predicting response early on during the treatment 94 95 course. 96 Funding: 97 Daiichi Sankyo Inc, Cancer Research UK, EU IMI2 IMMUCAN, UK Medical Research 98 Council, European Research Council (335326), National Institute for Health Research and The 99 Institute of Cancer Research. 100 101 102

#### 103 INTRODUCTION

Recurrent (R) or metastatic (M) Head and Neck Squamous Cell Carcinoma (HNSCC) is associated with a poor prognosis. Until the KEYNOTE-048 study was published in 2019(1), the standard-of-care, first line systemic treatment was the EXTREME regimen, consisting of a platinum-based chemotherapy regimen and cetuximab, an anti-EGFR monoclonal antibody(2). Even now, for patients with programmed death ligand 1 (PD-L1) negative tumors or those with contraindications to the use of anti-PD1 immunotherapy, the EXTREME regimen remains a first line standard-of-care.

While effective, this regimen is associated with toxicities. One of the key challenges for the treating physician is to identify the patients who would benefit from this treatment regimen. A predictive biomarker signature for patients with advanced HNSCC will help individualize discussions with patients regarding the risk-benefit balance of this treatment regimen and may guide patients who are likely to perform poorly towards alternative therapy regimens or clinical trials.

It is made available under a CC-BY 4.0 International license .

117 The absence of predictive biomarkers in this patient cohort represents a significant clinical 118 unmet need. Until the development of PD-L1 as a biomarker for immunotherapy, efforts to 119 generate biomarkers in HNSCC have focused on gene-expression profiles, which are 120 dependent on the availability of tumor tissue and are only performed on pre-treatment 121 samples(3, 4). Signatures based on a single biological modality and taken at a single timepoint 122 may be insufficient to predict outcomes, as response to therapy relies on a dynamic interplay 123 between cancer genomics, immune profile, tumor microenvironment, and clinicopathological 124 characteristics of the patient receiving treatment (5, 6).

Efforts to develop a machine learning model to stratify survival risk by combining genetic and clinicopathological characteristics have revealed some success in advanced oral squamous cell carcinoma(7). We hypothesize that a multimodal analysis, taking into account both clinicopathological and laboratory-based biological covariates at different timepoints, would provide better predictive value.

130 We prospectively collected peripheral blood samples from a Phase 2 trial in R/M HNSCC (NTC02633800)(8), which utilized a modified EXTREME regimen as a backbone, and 131 132 conducted a parallel exploratory analysis with the aim of generating a biomarker signature 133 which would predict outcomes to treatment. We hypothesized that the detailed definition of a 134 broad immune cell signature could contribute to the development of assays employing liquid 135 biopsies to predict clinical outcomes. We also incorporated the analysis of two circulating 136 microRNAs (miRNAs); miR-21-5p and miR-142-3p, which have previously demonstrated 137 prognostic and predictive utility(9, 10). As the trial investigated the efficacy of an anti-ErbB3 138 antibody, patritumab, administered alongside an anti-EGFR antibody, we simultaneously 139 analyzed EGFR-ErbB3 dimerization using Förster Resonance Energy Transfer (FRET) and 140 included it in our analysis.

It is made available under a CC-BY 4.0 International license .

By extracting information from patient samples at baseline and after the first cycle of treatment within this trial, we aimed to develop a multimodal predictive signature for the EXTREME regimen based on a novel Bayesian multivariate model. This can serve as a noninvasive risk stratification for patients with R/M HNSCC using only peripheral blood, guiding the clinician towards the likelihood of success early during the treatment course.

146

#### 147 MATERIALS AND METHODS

#### 148 Study Design

149 The clinical study design of the Phase 2 study (NCT02633800) and its associated 150 exploratory analysis are shown in Figure 1A. 87 patients were enrolled in the clinical trial. 151 Peripheral blood samples were collected before initiation of treatment (C1) and immediately 152 before the second cycle of treatment (C2). 31 patients were excluded due to incomplete paired 153 biological datasets, leaving 56 patients for analysis. Amongst these patients, there was no 154 difference in PFS as demonstrated by Kaplan-Meier survival curve analysis (Supplementary 155 Figure 1) regardless of whether the patients received patritumab, which reflected the results 156 published in the clinical trial. The baseline clinical characteristics of these 56 patients are shown 157 in Supplementary Table 1.

PBMC samples were analyzed using Flow Cytometry to generate unique immunological subpopulations. Exosomes were extracted from the serum and analyzed for EGFR-ErbB3 dimerization and miRNA-21-5p and miRNA-142-3p (Figure 1B). These analyses yielded a total of 29 unique biological covariates. Each covariate was obtained in pairs (C1 and C2), generating a total of 58 laboratory-based covariates for the multivariate analysis (Figure 1B). To mitigate individual baseline variations between patients, we evaluated changes between C1 and

It is made available under a CC-BY 4.0 International license .

164 C2 (in the form of log-fold change, LFC, of the variable of interest) instead of absolute values of 165 those parameters. A list of the laboratory-based and clinical covariates is provided in 166 Supplementary Table 2.

167 The baseline clinical characteristics, as well as value of the laboratory-based covariates at 168 baseline and after one cycle of treatment, did not significantly differ between the placebo and 169 patritumab cohorts (Supplementary Table 3). Therefore, in this exploratory analysis, samples 170 from both the control and investigational arms were analyzed together. The effect of adding the 171 investigational product, patritumab on Progression-Free Survival (PFS) was evaluated by 172 including it as an independent clinical covariate, denoted as 'Drug', in our multivariate analysis.

Written informed consent was obtained. Approval was obtained from ethics committees(Research Ethics Committee reference: 15/LO/1670).

It is made available under a CC-BY 4.0 International license .



- 176
- 177

Figure 1: Peripheral blood samples from the clinical trial were prospectively analyzed using a
multimodality platform (A) Schematic of clinical trial design and timepoints at which peripheral blood was obtained.
(B) Fifty-six (n=56) paired blood samples, obtained pre-treatment (C1) and after one cycle of treatment (C2) were
subjected to Flow Cytometry, FRET-FLIM imaging and ddPCR analysis.

It is made available under a CC-BY 4.0 International license .

#### 183 Statistical Analysis

184 To examine whether the various features and distribution of survival indices indicated 185 different prognostic outcomes, we built a model for predicting Progression-Free Survival (PFS). 186 Covariates were ranked by importance and selected by Bayesian multivariate proportional 187 hazards regression with backward elimination(11). We derived two models using separate 188 datasets: firstly, a baseline predictive model containing a dataset of 42 baseline covariates (29 189 laboratory parameters at baseline, C1, and 13 clinical characteristics). The second, a combined 190 predictive model, consists of 71 covariates, i.e. the 42 baseline covariates and a further 29 191 derived from the change in lab-based parameters between C1 and C2, measured by log fold 192 change (LFC) of the variable of interest.

193 The relative efficiency of the predictive model was assessed by using C-index (a metric 194 proposed by Harrell(12) to evaluate the accuracy of predictions made by an algorithm) and rank 195 correlation of the signature-generated risk scores with survival time. The number of significant 196 covariates in each prediction signature was determined with the aim of avoiding overfitting of the 197 signature to the study data using the "batch regression" option of the Saddle Point Signature 198 software (Saddle Point Science Ltd., London, UK), according to methods that were previously 199 published(13, 14). This is particularly important with small number of patients where an 200 independent test set is not possible. Systematic iterative covariate rejection and cross-validation 201 (5000 iterations) allowed for the selection of an optimal covariate set to avoid overfitting though 202 inclusion of too many covariates. The optimal set can be chosen in two ways, either based on 203 the peak prediction performance of cross-validation, or the more stringent method that equally 204 penalizes validation performance and overfitting (defined as the deviation between training and 205 validation performance). All signatures presented were chosen using the more stringent criterion 206 and data for all covariates is also presented for the purposes of identifying covariates that may

It is made available under a CC-BY 4.0 International license .

207 be important but do not quite meet the criterion. The regression included covariates 208 representing the missingness of data to account for the possibility that patient or sample 209 selection/rejection (for any reason) is biased with respect to outcome and therefore could be 210 informative. The missing data was imputed with the mean for that covariate. The importance 211 and significance of covariates can be judged by their assigned beta value ( $\beta$ ) in the proportional hazards model, and corresponding hazard ratio (HR) equal to  $e^{2\beta}$ . A negative beta value reflects 212 213 a lower risk of developing an event. The Signature software additionally judges the performance 214 of similar randomized data, which most often has beta values around zero and within a critical 215 range, such that any real covariate that has a beta value outside this critical range can be 216 judged to be performing significantly better than randomized data. This adds additional 217 confidence in the absence of an independent validation test set.

#### 218 Flow Cytometry

219 Frozen PBMC samples were thawed and stained with Fixable viability dye (Yellow 220 Live/Dead<sup>™</sup>, Fisher Scientific) followed by two different panels of membrane markers. A panel 221 for T cells included CD3, CD4, CD8, CD25, CD45RO, CD127, CCR7, and HLA-DR. A panel for 222 B cells and monocytes included CD3, CD19, CD24, CD38, CD27, IgD, CD33, CD11b, CD14 223 and CD16 (full list of both antibody panels in Supplementary Table 4). These two panels allow 224 definition of immune cell populations as described in Supplementary Figure 2. Patients' samples 225 and corresponding Fluorescence Minus One (FMO) Controls were acquired in a Fortessa II flow 226 cytometer (BD, Berkshire, UK) and analyzed with FlowJo software (Tree Star). Populations 227 were quantified as proportions of their respective parent population.

#### 228 Isolation of Serum exosomes

Exosomes were prepared using an optimized centrifugation method(15). Diluted serum was centrifuged at 300xg for 10 min to remove cell debris, 5000xg for 20 min to remove large vesicles and membrane fragments, and 12,200xg for 30 min to deplete microvesicles. This was followed by 100,000xg ultracentrifugation for 120 min at 4 °C to pellet exosomes with a TLA-55 rotor (Beckman Coulter). After a second 100,000xg ultracentrifugation for 60 minutes, the resulting pellets were washed and resuspended in PBS. Purified exosomal fractions were diluted and used for nanoparticle tracking analysis (NTA) using a Nanosight LM-14 system.

#### 236 RNA Extraction and miRNA Expression Analysis

237 RNA from cancer patients' serum exosomes was extracted using the TRIzol<sup>™</sup> Plus 238 RNA Purification Kit (Thermo Fisher, UK) according to the manufacturer' instructions. 239 Quantification of gene expression in circulating exosomes was performed by ddPCR (Bio-Rad 240 QX100 system). Normalization of the RNA, between cycle 1 and cycle 2 therapy of each patient, 241 was performed using the expression levels of the housekeeping gene 18S (Assay ID, 242 Hs99999901 s1). For each sample, equal volume of RNA was used as template and cDNA synthesis performed using the SuperScript® VILO™ MasterMix (Thermo Fisher, UK) according 243 to the manufacturer' instructions. MicroRNAs were reverse-transcribed individually using the 244 245 TaqMan<sup>™</sup> MicroRNA Reverse Transcription Kit (Thermo Fisher, UK). For each sample, 246 the normalized amount of RNA was reverse transcribed in a 15 µl reaction using the standard 247 protocol and primers specific for each miRNA: miR-21-5p (assay ID, 000397), miR-142-3p 248 (assay ID, 000464). Then 7.5  $\mu$ l of cDNA was added to a 20  $\mu$ l reaction containing 12.5  $\mu$ l 2X ddPCR Supermix for Probes (Bio-Rad) and 1 µl 20X TagMan miRNA PCR primer probe set; 249 each reaction was carried out in duplicate. Thermo cycling conditions were as it follows: 95 °C 250 251 for 10 min, then 50 cycles of 95 °C for 10 sec and 61 °C for 30 sec and a final inactivation step at 98 °C for 12 min. PCR products were analyzed using the QuantaSoft™ Software (Bio-Rad). 252

It is made available under a CC-BY 4.0 International license .

#### 253 ErbB3-EGFR Dimer Quantification in Exosomes

Exosomes were imaged on an 'Open' Fluorescence Lifetime Imaging Microscopy (FLIM) system (16). Analysis was performed with the TRI2 software (v2.7.8.9, CRUK/MRC Oxford Institute for Radiation Oncology, Oxford) as described previously (17, 18). Interfering effects of autofluorescence were minimized with a lifetime filtering algorithm and the FRET efficiency value for each patient calculated by:  $FRET=1-\frac{\tau_{DA}}{\tau_D}$ , where t<sub>D</sub> and t<sub>DA</sub> are the average lifetime of Alexa Fluor546 in the matching donor (D) and donor-acceptor (DA) images.

260

#### 261 Imaging Mass Cytometry

FFPE histological slides were stained with a panel of metal conjugated antibodies (full list of antibodies listed in Supplementary Table 5).

In brief, antigen retrieval was performed on a Ventana Bench Mark Ultra with CC1 buffer (Roche, 950-224). Slides were blocked for 1 hour at RT in 5% BSA, 5mg/ml human IgG in PBS and stained overnight at 4°C in 4% BSA, PBS. DNA counterstain was performed with Iridium (Fluidigm, 201192B) 125nM in PBS for 30 minutes at room temperature.

Ablation and data acquisition were performed on a Fluidigm Hyperion located within our Biomedical Research Centre. Imaging analysis was performed using the RUNIMC R package: RandomForest for classification and regression, Raster and SF for image manipulation and segmentation (see <u>https://www.biorxiv.org/content/10.1101/2021.09.14.460258v1</u>, with code available here: <u>https://github.com/luigidolcetti/RUNIMC</u>).

#### 273 RESULTS

The Model with Baseline Covariates Reveal Immune Subpopulations and Age PredictPFS

It is made available under a CC-BY 4.0 International license .

276 Bayesian multivariate proportional hazards regression was performed on the 42 277 covariates derived at baseline (C1) and PFS outcome. We utilized the stringent selection criteria 278 based on a proportional hazards regression model to minimize overfitting based on the cross-279 validation performance (Figure 2A). This revealed two baseline immune subpopulations with a 280 beta value which exceeded the critical beta-value threshold, ie CD14+CD16+CD33+CD11b+ 281 monocytes (thereafter referred to as CD33+CD14+ monocytes according to previous 282 nomenclature(19)) and double negative (CD27-IgD-) B cells (DN B cells), as well as one clinical 283 covariate – age (Figure 2B). Missingness covariates were included in this analysis and did not 284 affect the outcome of the signature.

Evaluation of the individual beta values reveal that baseline CD33+CD14+ monocytes and double negative B cells have a beta,  $\beta$  value of -1.05 and -0.53 respectively, and hence a higher baseline value of both populations is predictive of better PFS. Age, with a  $\beta$  value of 0.47, is associated with poorer PFS. The hazard ratios (HR) of the individual covariates are depicted in Figure 2C.

The risk scores generated from this signature were split at the median value to generate low-risk and high-risk cohorts (Figure 2D). The median overall survival of the low-risk and highrisk cohorts in this baseline predictive signature are 8.3 and 3.6 months respectively (log rank pvalue = 6.0 e-5) with a rank correlation of -0.29. The C-index of the predictive signature based completely on baseline parameters is 0.661. The risk score equation is given in Supplementary Figure 4.

It is made available under a CC-BY 4.0 International license .



297 Figure 2 High baseline CD33+CD14+ monocytes and Double Negative B cells Predict Overall Survival 298 Covariates were ranked for importance and selected by a proportional hazards regression model with cross-validation and backward 299 elimination. (A) Signature performance (defined as the classifier accuracy for predicting patient survival < or > median) as a function 300 of the number of active covariates as covariates are eliminated based on cross validation performance. (B) Heatmap of covariate 301 strength (beta value in the proportional hazards model) for the active covariates. Covariate elimination revealed three covariates in 302 the optimal signature - CD33+CD14+ monocytes, Double Negative B Cells, and Age. (C) Forest plot of the three covariates within 303 Progression Free Survival risk score with dotted line indicating the range, around 1, of typical random covariates. (D) Progression 304 Free Survival Risk Signature Performance, low risk score (n=27) and high risk score (n=26). Log rank P-value = 6.0 e-5, with 305 numbers at risk demonstrated under Kaplan-Meier curve. The multivariate analysis resulted in risk signatures that are linear 306 combinations of weighted covariates. Their ability to predict outcome is demonstrated with data split by signature value.

307

308

309

It is made available under a CC-BY 4.0 International license .

#### 311 Incorporating laboratory-based covariates after one cycle of treatment improves ability to

#### 312 predict PFS benefit

We subsequently evaluated if the incorporation of early laboratory-based changes into the signature improves its predictive ability. A separate predictive model incorporating 29 new covariates, ie changes in laboratory-based parameters between cycle 1 and cycle 2 was generated.

317 As before, we used a proportional hazards regression to determine a set of variables which 318 predict PFS. A total of six covariates were identified - three immune subpopulations with 319 negative beta values and hence are associated with better survival, ie baseline CD33+CD14+ 320 monocytes, baseline CD4 memory regulatory T cells (HLA-DR CD45RO<sup>+</sup>Tregs) and an increase 321 in CD8 effector memory T cells (CD45RO<sup>+</sup>CCR7<sup>-</sup>). An increase in two subpopulations, CD8 322 Central Memory T Cells (CD45RO<sup>+</sup>CCR7<sup>+</sup>) and CD3 T cells, was associated with inferior PFS. 323 The hypopharyngeal primary tumor site was also associated with a poorer PFS (Figure 3A and 324 B).

A multivariate analysis employing linear combinations of these six weighted covariates 325 326 generated a risk signature. Their ability to predict outcome is demonstrated with data split by 327 risk score, shown in Figure 3C. In this combined predictive signature, the median overall 328 survival of the low-risk and high-risk cohorts are 6.8 and 3.6 months respectively (log rank p-329 value 0.004) with a rank correlation of -0.38 (Figure D). The C-index of the predictive signature, 330 which incorporates baseline variables with changes in laboratory parameters after one cycle of 331 treatment, is 0.757. Both values are greater in the combined signature when compared to the 332 signature which only accounts for baseline values. The risk score equation is given in 333 Supplementary Figure 4.

It is made available under a CC-BY 4.0 International license .



#### 334

335 Figure 3 Model Incorporating Laboratory Changes After One Cycle of Treatment Exhibit Improved Predictive Value (A) 336 Heatmap of covariate strength (beta value in the proportional hazards model) for the active covariates. Proportional hazards 337 regression revealed five immune subpopulations in the optimal model - Baseline CD33+CD14+ monocytes, Baseline CD4 Memory 338 Regulatory T Cells, LFC of CD8 effector memory T cells, LFC of CD8 central memory T cells and LFC of CD3 T cells. The primary 339 tumor site of hypopharynx also featured in the signature. A negative beta value is associated with lower risk score and hence better 340 progression free survival. LFC = Log Fold Change (B) Forest plot of the three covariates within Progression Free Survival risk score 341 (C) Progression Free Survival Risk Signature Performance, low risk score (n=29) and high risk score (n=24). Log rank P-value = 342 0.004, with numbers at risk demonstrated under Kaplan-Meier curve.

343

It is made available under a CC-BY 4.0 International license .

#### 345 EGFR-ErbB3 FRET may contribute to predictive signature

346 While the combined predictive signature comprised predominantly of immunological 347 parameters, there is a suggestion that FRET difference may carry a degree of predictive value. In Figure 3A (4<sup>th</sup> covariate from the bottom), the difference in EGFR-ErbB3 FRET (FRET.delta) 348 349 was associated with a negative beta value which suggests a better PFS. However, the stringency that we have applied to optimal covariate selection means that this covariate fell 350 351 marginally short of featuring in the eventual predictive signature. Nonetheless, this is the first 352 time that this assay has been used within the context of a randomized controlled trial in 353 exosomes and the suggested predictive value of the dimer warrants some discussion.

Figure 4A displays intensity images and donor lifetime map of exosomes labelled with anti-EGFR and anti-ErbB3 antibodies, along with an accompanying schematic (Figure 4B).

By dividing the patients with available FRET values by the median FRET.delta (n=43), there was a suggestion that patients with a high FRET.delta exhibited a better PFS than patients with a low FRET.delta. This difference was not statistically significant (p=0.2), and the predictive capacity of this univariate is limited (Rank Correlation = -0.132, C-index = 0.561, Figure 4C). Nonetheless, these results suggest a trend within a small patient cohort and can be explored in future prospective studies.

It is made available under a CC-BY 4.0 International license .



362

363 Figure 4 FRET/FLIM fluorescence assay of circulating exosomes extracted from patients (A) Time-resolved fluorescence 364 intensity images and donor lifetime map of exosomes labelled with Anti-EGFR-IgG-Alexa 546 and Anti-ErbB3-IgG-Cy5 extracellular 365 antibodies (B) Schematic illustration of the fluorescent labelling geometry on exosomes and distance dependence of FRET 366 efficiency (C) Progression Free Survival of subpopulations divided by median FRET difference, FRET.delta low (n=21) and 367 FRET.delta high (n=22). Log rank P-value = 0.2, with numbers at risk demonstrated under Kaplan-Meier curve.

368

# Imaging Mass Cytometry of Tissue Reveals Correlation of CD33+CD14+ Myeloid Cell Subpopulation Between Tissue and Peripheral Blood

It is made available under a CC-BY 4.0 International license .

Having established that immune subsets in peripheral blood predict therapeutic response within a multivariate signature, we subsequently explored the relationship between the immune findings in peripheral blood with tumor infiltrating leukocytes (TILs). We obtained sufficient tissue from the biopsy at trial enrolment for in-depth profiling by imaging mass cytometry from four patients.

376 Standard FFPE samples from these four patients were processed as described previously, and 377 the results clustered in an unsupervised fashion. The range of PFS for these patients was 378 between 1.6-11.2 months (Figure 5A). Representative images of nuclear staining, overlaid with 379 pixel-level classification, are shown for the four patients in Figure 5B. Heatmaps representing 380 the distribution of cell phenotypes for each patient, as expressed by a two-phase classification 381 conducted at pixel and cell level, are illustrated in Figure 5C. The list of cell populations 382 characterized by imaging mass cytometry, alongside their detailed signature, is shown in 383 Supplementary Table 6.

A total of 7 cell clusters identified on tissue mass cytometry were significantly different between the four patients. CD33+CD14+ monocytes (cluster B-37) which featured in both baseline and combined signatures, exhibited a diminishing trend of abundance across patients 1 to 4. Having observed this trend, we subsequently correlated the levels of this population in tissue with blood. The proportion of CD33+CD14+ monocytes in the blood was 17.4%, 5.39%, 2.0% and 2.47% for patients 1-4 respectively, suggesting a meaningful concordance between the levels of this subpopulation in the patient tissue and peripheral blood.

It is made available under a CC-BY 4.0 International license .



It is made available under a CC-BY 4.0 International license .

394 Figure 5 Image Mass Cytometry of Four Patient Samples Using a 29 Marker Panel Analysis (A) Waterfall plot of 395 Progression Free Survival showing the four patients used for imaging mass cytometry analysis (B) Representative 396 image of the nuclear staining overlaid with pixel-level classification (C) Heatmaps representing the distribution of cell 397 phenotypes for each patient, as expressed by a two-phase clustering conducted at pixel level (columns) and at cell 398 level (rows). Data represent row-normalized areas. Red tiles, which represent hot spots of classification concordance, 399 are further described in panel D (D) Differential analysis of the classification hotspots presented in panel C, 400 highlighting cell populations which were significantly different between patients. Data are presented as sum of areas 401 of positive cells normalized to the total area of the ROI and expressed as µm<sup>2</sup> (x1000) per mm<sup>2</sup>. Statistics: \*adi-402 p<0.05, \*\*adj-p<0.01, \*\*\*\*adj-p<0.001, Pairwise Wilcoxon Rank Sum Tests with Benjamini correction, n≥3.

# 403 CD33+CD14+ Monocyte Population have high HLA-DR expression and univariate 404 predictive value

405 Due to the consistency with which the CD33+CD14+ monocyte population appeared across our 406 study, we wanted to further characterize this population to determine its phenotype. During the 407 process of drafting this manuscript, our lab was concurrently processing PBMCs from a cohort 408 of patients at risk of developing lung cancer. Using a second flow cytometry staining panel 409 which incorporates an alternative set of markers on these samples, we further characterized this 410 monocyte subpopulation, affirming that these CD14+CD16+CD33+CD11b monocytes also 411 express high levels of HLA-DR and CD11c (Supplementary Figure 3). This affirms that our 412 population of interest closely resembles the previously described CD33+CD14+ monocytes(19).

MiRNA signatures have been implicated as a useful classifier for myeloid cell subsets(20). By correlating the miRNA changes in our study with this monocytic subpopulation, a significant correlation was identified between the log fold changes of miR-21-5p with the corresponding log fold changes of CD33+CD14+ monocytes (Pearson's r=0.4343, p=0.02092, Figure 6A). We also investigated the potential of baseline CD33+CD14+ monocyte levels at predicting PFS. Figure 6B illustrates a Kaplan-Meier curve of PFS by median CD33+CD14+ monocyte level, generating

It is made available under a CC-BY 4.0 International license .

a modest split which surpassed conventional statistical significance (log rank p-value = 0.03.
However, the predictive capacity of CD33+CD14+ monocytes as a covariate was limited (Cindex 0.593, rank-correlation 0.22). These values were inferior in predictive capacity compared
to the baseline signature, which employed three covariates, and the combined predictive
signature which employed six covariates (Figure 6C).



;	Covariates	Number of	Concordance Index	Magnitude of	Log Rank
		Covariates		Rank Correlation	
	CD33+ Monocytes	1	0.593	0.22	0.03
	Baseline Signature	3	0.661	0.29	6e-05
	Combined Signature	6	0.757	0.40	0.004

#### 424

Figure 6 CD33+CD14+ Monocytes have some predictive value but predictive ability of signature is maximized with longitudinal sampling of peripheral blood (A) (B) Correlation between miRNA 21 fold change and CD33+ monocyte fold change after one cycle of treatment. Correlation efficient of 0.4343, p value 0.02092. miR21: microRNA-21-5p, lfc: log-fold change (B) Kaplan-Meier Curve of PFS split by median CD33+CD14+ value, log rank p-value 0.03 (C) Table summarises C-index, rank correlation and log-rank p value based on type and number of covariates. There is an increasing C-index and rank correlation upon the strict selection of additional covariates into the predictive signature, with the strongest signature incorporating six covariates combining values from baseline and after one cycle of treatment.

It is made available under a CC-BY 4.0 International license .

#### 432 **DISCUSSION**

There is a clinical unmet need to identify predictive biomarkers for treatment in head and neck cancer. Gene expression profiling has revealed promising initial results in this domain but have been limited to HPV positive HNSCC, which inherently have better prognoses. Recent developments in the field of immunotherapy in HNSCC have focused on tissue-based biomarkers, such as PD-L1, but when used in isolation these have not been sufficiently predictive at identifying patients who would benefit(1).

439 While uni-modal biomarkers may offer some predictive value, the biology of HNSCC and 440 likelihood of response to treatment is likely to be dictated by an interplay between tumor 441 immunity, genomic signatures and a host of clinicopathological characteristics. There is an 442 increased interest in peripheral blood biopsies in recent years, particularly in the context of 443 peripheral blood mononuclear cells(21). The use of peripheral blood in deriving biomarkers 444 mitigates a few limitations posed by tissue biopsies - particularly the accessibility and amount of 445 tissue required. The ease of obtaining liquid biopsies also facilitates longitudinal monitoring of 446 response to treatment.

447 To our knowledge, ours is the first piece of work integrating multiple biological covariates 448 derived from peripheral blood to generate a signature which predicts treatment response. We 449 also demonstrate the effectiveness of sequential monitoring of peripheral blood variables and 450 the advantage of longitudinal monitoring at enhancing prediction of response, as shown by the 451 combined predictive signature. By employing cross validation iterations to estimate training and 452 validation errors, implementing advanced overfitting correlation protocols, using built-in 453 corrections for informative data missingness, and probabilistic covariate removal, we were able 454 to derive a robust optimal covariate set which correlates with PFS. This combination of analyses 455 has been shown to produce robust signatures that do generalize to unseen data(13).

It is made available under a CC-BY 4.0 International license .

456 The biological components of the predicted model warrant discussion. The only clinical covariate to feature in the combined risk signature is the hypopharyngeal SCC sub-site, which 457 458 was adversely correlated with PFS. This corroborates previous findings that the 5-year relative 459 survival of patients with hypopharyngeal SCC is consistently the worst amongst different 460 anatomical HNSCC sub-sites (22, 23). The propensity of hypopharyngeal tumors to present at 461 the *de novo* advanced stage(24) and the density of submucosal lymphatics in this anatomical 462 region translates into these patients inherently performing worse - lending support to the robust 463 nature of our predictive signature. The notable absence of patritumab (denoted as 'Drug') in our 464 predictive signatures is also consistent with the outcome of the Phase 2 Clinical Trial where the 465 addition of this investigational medicinal product did not produce any benefit to PFS(8).

466 CD33+CD14+ monocytes demonstrated predictive capacity both as a univariate and as 467 a prominent feature in both baseline and combined signatures. Monocytes are a heterogenous 468 cell population, and phenotypic and functional characterization of monocyte subsets is a rapidly 469 Our identified emerging field(25). population of interest, 470 CD14+CD16+CD33+CD11b+CD11c+HLA-DR+ monocytes, intermediate resemble an 471 monocyte phenotype. This subset remains one of the most poorly characterized monocytic 472 subpopulations so far but have previously been linked to diverse immunological functions 473 including antigen processing and presentation, angiogenesis, and monocyte activation(26). 474 Interestingly, a significant correlation between the changes in miRNA-21-5p with changes in this 475 monocytic subpopulation was detected, supporting previous suggestions that miRNA signatures 476 can be a useful indicator of the functional state of myeloid cell subsets in cancer (20). The 477 predictive capacity of this subpopulation warrants investigation and further characterization in 478 future studies.

479 Beyond the interest in individual covariates, our study also reveals the potential of using 480 liquid-based biological outcomes to predict outcome to therapy. It has been widely recognized

It is made available under a CC-BY 4.0 International license .

that even the most utilized biomarkers, such as tumor PD-L1, have limited predictive value when used in isolation. Our study reveals that a targeted multimodality signature, obtained through longitudinal sampling of peripheral blood, is able to augment this predictive capacity and better identify patients who would benefit from a particular treatment regimen.

485 There are a few limitations to the study. The absence of overall survival (OS) within our 486 current dataset represents one of the shortcomings of the study and it would have been 487 interesting to assess whether the immune markers, particularly the CD33+CD14+ monocytic 488 population, predict survival in the longer term. However, the accuracy of the predictive signature 489 for OS would have been diluted by a variety of subsequent treatment regimens. Secondly, due 490 to tissue scarcity, we only managed to obtain sufficient biopsy tissue from four patients for in-491 depth profiling by mass cytometry to investigate the correlation of the monocytic population 492 between tissue and blood.

The present study shows that the combination of biomarkers established prospectively by liquid biopsies early in the treatment course offers potential for the provision of personalized treatments to patients (27). The post-stratification survival curves in our study demonstrate markedly different progression free survivals as a testament to this robust statistical model, and could represent an invaluable guide to clinicians during the initial stages of treatment.

498 FUNDING

499 This work was supported by a grant from Daiichi Sankyo Inc ('Identification of Non-

500 Invasive Treatment Stratification and Longitudinal Monitoring Markers for Patritumab/Cetuximab

501 Combination Therapy'). This work was also supported by Cancer Research UK funding support

502 to King's College London – UCL Comprehensive Cancer Imaging Centre (CR-UK & EPSRC),

503 Cancer Research UK King's Health Partners Centre at King's College London, and Cancer

504 Research UK UCL Centre; University College London (PRB) – Early Detection Award

505 (C7675/A29313); as well as CRUK City of London Centre. MG and KN are supported by Cancer
506 Research UK Clinical Training Fellowships (Award number 163011 for MG and 176885 for KN).
507 LD is supported by EU IMI2 IMMUCAN (Grant agreement number 821558). GA and JMV are
508 supported by CRUK Early Detection and Diagnosis Committee Project grant.

509 JWO is supported by the UK Medical Research Council (MR/N013700/1) and is a KCL 510 member of the MRC Doctoral Training Partnership in Biomedical Science. FW is also supported 511 by the UK Medical Research Council (MR/N013700/1). JNA is funded by a grant from the 512 European Research Council (335326).

513 MDF supported the UCL/UCLH NIHR Biomedical Research is by 514 Centre and runs early phase studies in the NIHR UCLH Clinical Research Facility supported by the UCL ECMC, MTD and KH acknowledge funding support from The 515 516 Institute of Cancer Research/Royal Marsden Hospital NIHR Biomedical Research Centre.

517 This research was funded/supported by the National Institute for Health Research 518 (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and 519 King's College London and/or the NIHR Clinical Research Facility. The views expressed are 520 those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of 521 Health and Social Care.

# 522 DATA AVAILABILITY STATEMENT

523 The data generated in this study and used for multivariate modelling are available from the UCL
524 repository: <u>https://doi.org/10.5522/04/16566207.v1</u>

- 525 **NOTES**
- 526 Affiliations of authors

#### It is made available under a CC-BY 4.0 International license .

527 Breast Cancer Now Research Unit, School of Cancer & Pharmaceutical Sciences, King's 528 College London, London, UK (FF-B,TN); Richard Dimbleby Laboratory of Cancer Research, 529 School of Cancer & Pharmaceutical Sciences, King's College London, London, UK 530 (GA,GW,FW,RM,TN); Comprehensive Cancer Centre, School of Cancer & Pharmaceutical 531 Sciences, King's College London, London, UK (PRB); UCL Cancer Institute, Paul O'Gorman 532 Building, University College London, London, UK (PRB, JMV, MG, KN, JB, MDF, TN); Tumor 533 Immunology Group, School of Cancer & Pharmaceutical Sciences, King's College London, 534 London, UK (JWO, JNA), Systems Cancer Immunology, School of Cancer & Pharmaceutical 535 Sciences, King's College London, London, UK (SK), Daiichi Sankyo Incorporated, New Jersey, USA (JD,JG); The Institute of Cancer Research/The Royal Marsden Hospital National Institute 536 537 for Health Research Biomedical Research Centre, London, UK (KH, MTD); Institute for 538 Mathematical and Molecular Biomedicine, King's College London, UK (ACC), Saddle Point Science Ltd, London, UK (ACC), Centre for Immunobiology and Regenerative Medicine, Barts & 539 540 The London School of Medicine and Dentistry, Queen Mary University of London, London, UK 541 (Current Affiliation for FF-B); Barts Health NHS Trust, London, UK (Current Affiliation for MG).

#### 542 Conflict of Interests

543 KN has received honoraria from Pfizer, GSK/Tesaro and Boheringer Ingleheim, and has had travel/accommodation/expenses paid for by Tesaro. MDF has received institutional 544 545 research funding from AstraZeneca, Boehringer-Ingelheim, Merck and MSD and serves in a consulting or advisory role to Achilles, Astrazeneca, Bayer, Bristol-Myers Squibb, Celgene, 546 547 Guardant Health, Merck, MSD, Nanobiotix, Novartis, Oxford VacMedix, Pfizer, Roche, Takeda, 548 UltraHuman. KH has received honoraria from Amgen; Arch Oncology; AstraZeneca; 549 Boehringer-Ingelheim; Bristol-Myers Squibb; Codiak; Inzen; Merck; MSD; Pfizer; Replimune and 550 is on a speakers' bureau for Amgen, AstraZeneca; Bristol-Myers Squibb; Merck, MSD; Pfizer.

551 KH has also received research funding from AstraZeneca, Boehringer-Ingelheim, MSD and552 Replimune.

JG and JD are both in employment with Daiichi Sankyo, and have stock and other 553 554 ownership interests, research funding within Daichii Sanyko and have had 555 travel/accommodation/expenses paid for by Daichii Sankyo. In addition, JD has also had stock and other ownership interests with Pfizer and received research funding from Novartis. ACCC 556 557 has stock and other ownership interests with Saddle Point Science Limited.

558 FW was initially funded by Daichii Sankyo as research assistant to conduct laboratory 559 work in the context of the translational aspect of this trial. SK has received research funding in 560 the form of a grant from Novartis and Celgene. TN has received research funding from 561 Astrazeneca and Daichii Sakyo. TN is a founder and shareholder in Nano Clinical Ltd, and PRB 562 is a shareholder.

563 FF-B,GA, GW,JMV,MG,JWO,JB and RM declare no conflicts of interests.

564 The funders had no role in the study design, the collection, analysis and interpretation of the565 data, the writing of the manuscript, and the decision to publish.

#### 566 ETHICS AND PERMISSIONS

- 567 Written informed consent was obtained for all patients who participated in the Phase 2 clinical
- trial. Approval was obtained from ethics committees (Research Ethics Committee reference:
- 569 15/LO/1670). Approval to procure and process a separate cohort of blood samples from patients
- 570 at risk of developing lung cancer was also obtained (IRAS ID: 261766).

# 571 ACKNOWLEDGEMENTS

- 572 We thank the patients who participated in this Phase 2 trial and the staff members at the study
- 573 sites who cared for them. We also thank Dr James Barrett for his assistance towards designing
- 574 the statistical techniques applied in this study.

### 575 PRIOR PRESENTATIONS

- 576 This study was presented at the 2018 American Society of Clinical Oncology (ASCO) Annual
- 577 Meeting in Chicago, June 1-5, 2018.

# 578 LIST OF SUPPLEMENTARY MATERIALS

- 579 Supplementary Figure 1: Kaplan Meier Curve of Progression Free Survival in Study Cohort
- 580 Supplementary Table 1: Demographic and Laboratory-Based Values of Patients Within Study
- 581 Supplementary Table 2: List of covariates entered into Bayesian Multivariate Analysis
- 582
- 583 Supplementary Table 3: Demographic and Laboratory-Based Values of Patients Separated by
- 584 Arm of Treatment Within Trial
- 585 Supplementary Figure 2: Gating Strategies For Definition of Peripheral Blood Immune
- 586 Populations
- 587 Supplementary Table 4: List of antibodies used in T cell panel and B cell-Monocyte Panel for
- 588 Flow Cytometry
- 589 Supplementary Table 5: List of Antibodies Used in Imaging Mass Cytometry (CyTOF) Analysis
- 590 for Definition of Immune Subpopulations
- 591 Supplementary Table 6: List of populations in imaging mass cytometry of tissue and
- 592 corresponding signature
- 593 Supplementary Figure 3: Gating strategy for further characterization of CD33+CD14+ monocytic
- 594 population using new patient cohort

It is made available under a CC-BY 4.0 International license .

#### 595 **REFERENCES**

596 Burtness B, Harrington KJ, Greil R, Soulieres D, Tahara M, de Castro G, Jr., et al. Pembrolizumab alone or 1. 597 with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the 598 head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. Lancet. 2019;394(10212):1915-28. 599 Van Cutsem E, Kohne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A, et al. Cetuximab and 2. 600 chemotherapy as initial treatment for metastatic colorectal cancer. N Engl J Med. 2009;360(14):1408-17. 601 3. Bossi P, Bergamini C, Siano M, Cossu Rocca M, Sponghini AP, Favales F, et al. Functional Genomics 602 Uncover the Biology behind the Responsiveness of Head and Neck Squamous Cell Cancer Patients to Cetuximab. 603 Clin Cancer Res. 2016;22(15):3961-70. 604 4. You GR, Cheng AJ, Lee LY, Huang YC, Liu H, Chen YJ, et al. Prognostic signature associated with 605 radioresistance in head and neck cancer via transcriptomic and bioinformatic analyses. BMC Cancer. 2019;19(1):64. 606 Cheerla A, Gevaert O. Deep learning with multimodal representation for pancancer prognosis prediction. 5. 607 Bioinformatics. 2019;35(14):i446-i54. 608 Huang Z, Zhan X, Xiang S, Johnson TS, Helm B, Yu CY, et al. SALMON: Survival Analysis Learning 6. 609 With Multi-Omics Neural Networks on Breast Cancer. Front Genet. 2019;10:166. 610 7. Tseng YJ, Wang HY, Lin TW, Lu JJ, Hsieh CH, Liao CT. Development of a Machine Learning Model for 611 Survival Risk Stratification of Patients With Advanced Oral Cancer. JAMA Netw Open. 2020;3(8):e2011768. 612 8. Forster MD, Dillon MT, Kocsis J, Remenar E, Pajkos G, Rolland F, et al. Patritumab or placebo, with 613 cetuximab plus platinum therapy in recurrent or metastatic squamous cell carcinoma of the head and neck: A 614 randomised phase II study. Eur J Cancer. 2019;123:36-47. Summerer I, Unger K, Braselmann H, Schuettrumpf L, Maihoefer C, Baumeister P, et al. Circulating 615 9. 616 microRNAs as prognostic therapy biomarkers in head and neck cancer patients. Br J Cancer. 2015;113(1):76-82. 617 Vahabi M BG, Di Agostino, S. MicroRNAs in head and neck squamous cell carcinoma: a possible 10. 618 challenge as biomarkers, determinants for the choice of therapy and targets for personalized molecular therapies. 619 Transl Cancer Res. 2021;10(6):3090-110. 620 Grigoriadis A, Gazinska P, Pai T, Irhsad S, Wu Y, Millis R, et al. Histological scoring of immune and 11. 621 stromal features in breast and axillary lymph nodes is prognostic for distant metastasis in lymph node-positive breast 622 cancers. J Pathol Clin Res. 2018;4(1):39-54. 623 12. Harrell FE, Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved 624 prognostic prediction. Stat Med. 1984;3(2):143-52. 625 Barber PR, Weitsman G, Lawler K, Barrett JE, Rowley M, Rodriguez-Justo M, et al. HER2-HER3 13. 626 Heterodimer Quantification by FRET-FLIM and Patient Subclass Analysis of the COIN Colorectal Trial. J Natl 627 Cancer Inst. 2020;112(9):944-54. 628 Shalabi A, Inoue M, Watkins J, De Rinaldis E, Coolen AC. Bayesian clinical classification from high-14. 629 dimensional data: Signatures versus variability. Stat Methods Med Res. 2018;27(2):336-51. 630 Monypenny J, Milewicz H, Flores-Borja F, Weitsman G, Cheung A, Chowdhury R, et al. ALIX Regulates 15. 631 Tumor-Mediated Immunosuppression by Controlling EGFR Activity and PD-L1 Presentation. Cell Rep. 632 2018;24(3):630-41. 633 16. Barber PR, Tullis ID, Pierce GP, Newman RG, Prentice J, Rowley MI, et al. The Gray Institute 'open' high-634 content, fluorescence lifetime microscopes. J Microsc. 2013;251(2):154-67. Barber PR, Ameer-Beg SM, Gilbey J, Carlin LM, Keppler M, Ng TC, et al. Multiphoton time-domain 635 17. 636 FLIM: Practical application to protein-protein interactions using global analysis. J R Soc Interface. 2009;6:S93-637 S105. 638 18. Rowley MI, Coolen ACC, Vojnovic B, Barber PR. Robust Bayesian Fluorescence Lifetime Estimation, 639 Decay Model Selection and Instrument Response Determination for Low-Intensity FLIM Imaging. PLOS ONE. 640 2016;11(6):e0158404. 641 Cravens PD, Hayashida K, Davis LS, Nanki T, Lipsky PE. Human peripheral blood dendritic cells and 19. 642 monocyte subsets display similar chemokine receptor expression profiles with differential migratory responses. 643 Scand J Immunol. 2007;65(6):514-24. 644 Bronte V, Brandau S, Chen SH, Colombo MP, Frey AB, Greten TF, et al. Recommendations for myeloid-20. 645 derived suppressor cell nomenclature and characterization standards. Nat Commun. 2016;7:12150.

It is made available under a CC-BY 4.0 International license.

- 646 21. Nixon AB, Schalper KA, Jacobs I, Potluri S, Wang IM, Fleener C. Peripheral immune-based biomarkers in 647 cancer immunotherapy: can we realize their predictive potential? J Immunother Cancer. 2019;7(1):325.
- 648 Gatta G. Capocaccia R, Botta L, Mallone S, De Angelis R, Ardanaz E, et al. Burden and centralised 22.
- 649 treatment in Europe of rare tumours: results of RARECAREnet-a population-based study. Lancet Oncol. 650 2017;18(8):1022-39.
- 651 Machiels JP, Rene Leemans C, Golusinski W, Grau C, Licitra L, Gregoire V, et al. Squamous cell 23.
- 652 carcinoma of the oral cavity, larynx, oropharynx and hypopharynx: EHNS-ESMO-ESTRO Clinical Practice 653 Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2020;31(11):1462-75.
- 654 Cadoni G, Giraldi L, Petrelli L, Pandolfini M, Giuliani M, Paludetti G, et al. Prognostic factors in head and 24. 655 neck cancer: a 10-year retrospective analysis in a single-institution in Italy. Acta Otorhinolaryngol Ital. 656 2017;37(6):458-66.
- 657 Ginhoux F, Mildner A, Gautier EL, Schlitzer A, Jakubzick C, Varol C, et al. Editorial: Monocyte 25. 658 Heterogeneity and Function. Front Immunol. 2020;11:626725.
- 659 Zawada AM, Rogacev KS, Rotter B, Winter P, Marell RR, Fliser D, et al. SuperSAGE evidence for 26. 660 CD14++CD16+ monocytes as a third monocyte subset. Blood. 2011;118(12):e50-61.
- 661 27. Nenclares P, Gunn L, Soliman H, Bover M, Trinh A, Leslie I, et al. On-treatment immune prognostic score 662 for patients with relapsed and/or metastatic head and neck squamous cell carcinoma treated with immunotherapy. J
- 663 Immunother Cancer. 2021;9(6).
- 664



# Supplementary Figure 1. Kaplan Meier Curve of Progression-Free Survival in Study Cohort

Progression-Free Survival data was obtained from 53 patients, and stratified according to patients who received patritumab and the cohort which received placebo.

Overall se.

It is made available under a CC-BY 4.0 In	ternational licer
Age (mean (SD))	58.46 (8.78)
Drug = Placebo (%) Sex = M (%)	27 (48.2)
Site (%)	40 (82.1)
Aryepiglottis	1 ( 1.8)
Hypopharynx	12 (21.4)
Oral Cavity	20 (35.7)
Oropharynx	10 (17.9)
Other	5 ( 8.9)
Prev.Cetux = Yes (%) HPV Status = Pos (%)	1 (1.8)
Hereg.Status = Low (%)	27 (48.2)
Smoking (%)	
Current	18 (32.1)
Never	3 ( 5.4)
Other	1(1.8)
Dendritic.CD33.c1 (mean (SD))	4.49 (4.46)
T.cell.CD4.c1 (mean (SD))	50.07 (18.43)
T.cell.CD4.Cent.Mem.c1 (mean (SD))	27.99 (9.58)
T.cell.CD4.Effector.c1 (mean (SD))	9.15 (18.07)
CD4 Naive c1 (mean (SD))	31.77 (17.04)
CD4.Tregs.c1 (mean (SD))	8.97 (5.24)
CD4.Memory.Tregs.c1 (mean (SD))	62.15 (12.50)
CD4 Naive Tregs c1 (mean (SD))	22.45 (12.25)
T.cell.CD8.c1 (mean (SD))	35.27 (12.98)
CD8.Cent.Mem.c1 (mean (SD))	9.34 (4.87)
CD8.Effector.c1 (mean (SD))	38.74 (15.99)
CD8.Naive.c1 (mean (SD))	18.54 (13.92)
CD4.CD8.ratio.c1 (mean (SD))	1.85 (1.23)
B.cell.CD19c1 (mean (SD))	17.15 (8.62)
B.cell.meture.c1 (mean (SD)) B.cell.memory.c1 (mean (SD))	21.18 (12.04)
B.cell.Transitional.c1 (mean (SD))	7.30 (8.30)
B.cell.plasma.c1 (mean (SD))	0.36 (0.62)
B.cell.DN.c1 (mean (SD)) B.cell paive c1 (mean (SD))	13.84 (9.01)
B.cell.pre.switch.c1 (mean (SD))	4.82 (5.52)
B.cell.post.switch.c1 (mean (SD))	13.26 (8.83)
FRET.c1 (mean (SD))	0.00 (0.05)
normalised.miR.142.3p.c1 (mean (SD))	0.00 (0.00)
FRET.c1.missing (mean (SD))	0.14 (0.35)
miR.c1.missing (mean (SD))	0.11 (0.31)
Dendritic.CD33.lfc (mean (SD))	0.31 (1.29)
T.cell.CD3.lfc (mean (SD))	0.28 (0.57)
T.cell.CD4.lfc (mean (SD))	0.12 (0.63)
T.cell.CD4.Effector.lfc (mean (SD))	-0.37 (1.00)
T.cell.CD4.Effect.Mem.lfc (mean (SD))	0.01 (0.84)
CD4.Naive.lfc (mean (SD))	0.08 (0.98)
CD4. Memory, Tregs. Ifc (mean (SD))	-0.02 (0.30)
CD4.Mem.Act.Tregs.Ifc (mean (SD))	0.19 (0.77)
CD4.Naive.Tregs.lfc (mean (SD))	-0.10 (0.97)
T.cell.CD8.lfc (mean (SD)) CD8 Cent Mem lfc (mean (SD))	0.11 (0.41)
CD8.Effector.lfc (mean (SD))	-0.17 (0.62)
CD8.Effect.Mem.Ifc (mean (SD))	0.03 (0.33)
CD8.Naive.lfc (mean (SD))	-0.15 (0.85)
B.cell.CD19lfc (mean (SD))	-0.07 (0.89)
B.cell.mature.lfc (mean (SD))	0.03 (0.45)
B.cell.memory.lfc (mean (SD))	0.10 (0.45)
B.cell.plasma.lfc (mean (SD))	0.12 (2.03)
B.cell.DN.lfc (mean (SD))	-0.18 (0.61)
B.cell.naive.lfc (mean (SD))	-0.01 (0.21)
B.cell.post.switch.lfc (mean (SD))	0.14 (0.66)
normalised.miR.21.5.lfc (mean (SD))	0.25 (2.12)
normalised.miR.142.3p.lfc (mean (SD))	0.13 (1.98)
FRET.delta (mean (SD))	0.00 (0.05)
miR.lfc.missing (mean (SD))	0.11 (0.31)
immune.lfc.missing (mean (SD))	0.46 (0.50)
BOR.RECIST (mean (SD))	2.74 (0.68)
t.PFS (mean (SD))	5.48 (3.48)

## Supplementary Table 1. Demographic and Laboratory-Based Values of patients Group of 56

patients with HNSCC was analyzed and their characteristics indicated as above.

	Baseline Clinical			
PBMC-Based		Exosom	e-Based	Covariates (n=13)
Panel A	Panel B	FRET value	miRNA	
CD3 T cells	CD19 B cells	FRET-FLIM	miR 21.5	Patritumab (Drug)
CD4 T cells	Mature B cells	EGFR-ErbB3	miR 142.3p	Age
CD8 T cells	Transitional B cells			Sex
CD4 Central	Plasma B cells			Smoking Status
Memory	Double Negative B			Site:
CD4 Effector	cells			-Oral Cavity
Memory	Naïve B cells			-Oropharynx
CD4 Naïve	Pre switch B cells			-Hypopharynx
CD4 Tregs	Post switch B cells			-Aryepiglottis
CD4 Memory	CD33+CD14+			-Larynx
Tregs	Monocytes			-Other
CD4 Mem Act				Previous Cetuximab
Tregs				HPV Status: +ve/-ve
CD4 Naïve Tregs				Heregulin Status
CD8 Central				
Memory				
CD8 Effector				
CD8 Effector				
Memory				
CD8 Naïve				
CD4.CD8 ratio				

**Supplementary Table 2:** List of covariates entered into Bayesian multivariate analysis. 29 laboratory-based covariates were used (obtained at baseline and after one cycle of treatment), and combined with 13 baseline clinical covariates.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.22.21263942; this version posted September 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a  $CC-BY \neq 0$  International license.

It is ma	de available under a C	Patritumab	Placebo	pation
it is ma		58.17(8.59)	58.78 (9.14)	0.799
	Drug = Placebo (%)	0 ( 0.0)	27 (100.0)	<0.001
	Sex = M (%)	22 (75.9)	24 ( 88.9)	0.356
	Site (%)	4 ( 3 4)	0 ( 0 0)	0.498
	Hypopharynx	7 (24.1)	5 (18.5)	
	Larynx	3 (10.3)	5 (18.5)	
	Oral Cavity	12 (41.4)	8 ( 29.6)	
	Oropharynx	3 (10.3)	7 (25.9)	
	Prev.Cetux = Yes (%)	0(0.0)	2(7.4)	0.971
	HPV.Status = Pos (%)	4 (13.8)	4 (14.8)	1
	Hereg.Status = Low (%)	14 (48.3)	13 ( 48.1)	1
	Smoking (%)	7 (24.1)	11 ( 40 7)	0.36
	Ex	20 (69.0)	14 ( 51.9)	
	Never	1(3.4)	2 (7.4)	
	Other	1(3.4)	0(0.0)	
	Dendritic.CD33.c1 (mean (SD))	4.17 (4.71)	4.73 (4.38)	0.716
	T.cell.CD4.c1 (mean (SD))	49.68 (17.43)	50.36 (19.60)	0.919
	T.cell.CD4.Cent.Mem.c1 (mean (SD))	29.86 (10.13)	26.61 (9.18)	0.343
	T.cell.CD4.Effector.c1 (mean (SD))	17.41 (25.54)	3.07 (4.19)	0.022
	T.cell.CD4.Effect.Mem.c1 (mean (SD))	31.04 (17.90)	32.32 (16.85)	0.834
	CD4.Tregs.c1 (mean (SD))	8.79 (4.05)	9.11 (6.07)	0.864
	CD4.Memory.Tregs.c1 (mean (SD))	60.04 (14.54)	63.69 (10.91)	0.415
	CD4.Mem.Act.Tregs.c1 (mean (SD))	20.58 (11.36)	23.83 (12.99)	0.46
	CD4.Naive.Tregs.c1 (mean (SD))	18.84 (19.33)	11.62 (7.69)	0.148
	CD8.Cent.Mem.c1 (mean (SD))	8.83 (3.60)	9.74 (5.73)	0.619
	CD8.Effector.c1 (mean (SD))	42.45 (18.41)	35.90 (13.76)	0.273
	CD8.Effect.Mem.c1 (mean (SD))	30.91 (15.34)	35.26 (9.02)	0.338
	CD8.Naive.c1 (mean (SD))	17.81 (13.23)	19.10 (14.80)	0.807
	B.cell.CD19c1 (mean (SD))	16.69 (8.88)	17.50 (8.63)	0.789
	B.cell.mature.c1 (mean (SD))	55.65 (16.19)	60.64 (10.92)	0.284
	B.cell.memory.c1 (mean (SD))	22.00 (15.59)	20.56 (8.90)	0.731
	B.cell.Iransitional.c1 (mean (SD))	8.55 (12.20)	6.37 (3.37)	0.448
	B.cell.DN.c1 (mean (SD))	12.40 (7.87)	14.92 (9.83)	0.421
	B.cell.naive.c1 (mean (SD))	70.15 (15.33)	66.51 (14.13)	0.472
	B.cell.pre.switch.c1 (mean (SD))	5.71 (7.51)	4.16 (3.46)	0.421
	FRET.c1 (mean (SD))	-0.01 (0.05)	0.00 (0.06)	0.384
	normalised.miR.21.5.c1 (mean (SD))	0.00 (0.00)	0.01 (0.02)	0.531
	normalised.miR.142.3p.c1 (mean (SD))	0.00 (0.00)	0.00 (0.00)	0.687
	FRET.c1.missing (mean (SD))	0.07 (0.26)	0.22 (0.42)	0.105
	immune.c1.missing (mean (SD))	0.48(0.51)	0.11 (0.32)	0.928
	Dendritic.CD33.lfc (mean (SD))	0.14(1.02)	0.42 (1.47)	0.569
	T.cell.CD3.lfc (mean (SD))	0.25 (0.61)	0.30 (0.56)	0.833
	T.cell.CD4.ltc (mean (SD)) T.cell.CD4.Cent Mem.ltc (mean (SD))	0.08 (0.54)	0.15 (0.70)	0.796
	T.cell.CD4.Effector.lfc (mean (SD))	-0.53 (0.96)	-0.25 (1.03)	0.459
	T.cell.CD4.Effect.Mem.lfc (mean (SD))	-0.02 (0.86)	0.04 (0.86)	0.855
	CD4.Naive.lfc (mean (SD))	0.19 (1.07)	0.01 (0.94)	0.639
	CD4. Iregs.ifc (mean (SD)) CD4. Memory, Tregs.ifc (mean (SD))	-0.23 (0.91)	-0.19 (0.58)	0.0891
	CD4.Mem.Act.Tregs.Ifc (mean (SD))	0.39 (0.83)	0.06 (0.71)	0.255
	CD4.Naive.Tregs.lfc (mean (SD))	-0.28 (0.76)	0.02 (1.08)	0.432
	T.cell.CD8.lfc (mean (SD))	0.14 (0.38)	0.09 (0.44)	0.756
	CD8.Effector.lfc (mean (SD))	-0.34 (0.64)	-0.06 (0.60)	0.263
	CD8.Effect.Mem.lfc (mean (SD))	0.15 (0.33)	-0.05 (0.31)	0.125
	CD8.Naive.lfc (mean (SD))	0.02 (0.56)	-0.27 (1.00)	0.39
	CD4.CD8.ratio.lfc (mean (SD))	-0.19 (0.66)	-0.08 (0.89)	0.734
	B.cell.mature.lfc (mean (SD))	0.21(0.38)	-0.08 (0.46)	0.078
	B.cell.memory.lfc (mean (SD))	0.01 (0.46)	0.15 (0.45)	0.411
	B.cell.Transitional.lfc (mean (SD))	-0.21 (0.74)	-0.29 (0.80)	0.775
	B.cell.plasma.lfc (mean (SD))	0.43 (1.97)	-0.04 (2.10)	0.564
	B.cell.naive.lfc (mean (SD))	0.06 (0.23)	-0.06 (0.19)	0.126
	B.cell.pre.switch.lfc (mean (SD))	0.19 (0.82)	0.10 (0.56)	0.729
	B.cell.post.switch.lfc (mean (SD))	-0.09 (0.57)	0.08 (0.65)	0.459
	normalised.miR.21.5.lfc (mean (SD))	0.51(1.92)	0.13 (2.16)	0.734
	FRET.delta (mean (SD))	0.00 (0.05)	0.00 (0.05)	0.631
	FRET.delta.missing (mean (SD))	0.14 (0.35)	0.22 (0.42)	0.42
	miR.lfc.missing (mean (SD))	0.10(0.31)	0.11 (0.32)	0.928
	IMMUNE.Itc.missing (mean (SD)) BOR.RECIST (mean (SD))	0.59 (0.50) 2.64 (0.68)	0.33 (0.48)	0.06
	event.PFS (mean (SD))	0.71 (0.46)	0.76 (0.44)	0.713
	t.PFS (mean (SD))	4.94 (2.83)	6.08 (4.06)	0.237

Supplementary Table 3. Demographic and Laboratory-Based Values of patients separated by arm of

treatment on trial (placebo vs patritumab)



Supplementary Figure 2. Gating strategies for definition of peripheral blood immune populations. Peripheral immune cell populations analysed in this study are indicated in corresponding gates A) Subpopulations of  $CD3^+CD4^+$  and  $CD3^+CD8^+$  T cells were further identified by the use of markers CCR7 and CD45RO as effector (Eff., CD45RO CCR7), naïve (CD45RO CCR7<sup>+</sup>), central memory (CM, CD45RO<sup>+</sup>CCR7<sup>+</sup>) and effector memory (EM, CD45RO<sup>+</sup>CCR7). The activation status of CD4<sup>+</sup> regulatory T cells (Tregs, CD127<sup>-/10</sup> CD25<sup>hi</sup>) was defined by markers CD45RO and HLA-DR as indicated: naïve Tregs (HLA-DR CD45RO<sup>-</sup>), activated Tregs (Act. Tregs, HLA-DR<sup>+</sup>CD45RO<sup>+</sup>), and memory Tregs (Mem. Tregs, HLA-DR<sup>-</sup>CD45RO<sup>+</sup>). B) CD19<sup>+</sup> B cell subpopulations were identified with markers CD24, CD38 and CD27: plasma cells (CD38<sup>+</sup>CD27<sup>+</sup>); mature (CD24<sup>+</sup>CD38<sup>hi</sup>), memory (CD24<sup>+</sup>CD38<sup>-</sup>), transitional (TBC, CD24<sup>+</sup>CD38<sup>hi</sup>); or a combination of IgD and CD27:naïve (IgD<sup>+</sup>CD27<sup>-</sup>), memory pre-switch (IgD<sup>+</sup>CD27<sup>+</sup>), memory post-switch (IgD<sup>-</sup>CD27<sup>+</sup>),

or double-negative B cells (DN, mgD code). CD33<sup>+</sup>CD14<sup>+</sup> Moliocyles were defined as CD3 CD19<sup>+</sup> CD33<sup>+</sup>CD16<sup>+</sup>CD14<sup>+</sup>.

T cell Panel			B cell-Monocyte Panel			
Target	Clone	Company	Target	Clone	Company	
CD3	UCHT1	BD Biosciences	CD3	UCHT1	BD Biosciences	
CD4	SK3	BD Biosciences	CD11b	ICRF44	Biolegend	
CD8	RPA-T8	BD Biosciences	CD14	HCD14	Biolegend	
CD25	M-A251	BD Biosciences	CD16	3G8	BD Biosciences	
CD45RO	UCHL1	BD Biosciences	CD19	HIB19	Biolegend	
CD127	HIL-7R-M21	BD Biosciences	CD24	ML5	Biolegend	
CCR7	GO43H7	Biolegend	CD27	LI28	Biolegend	
HLA-DR	G46-6	BD Biosciences	CD33	P67.6	Biolegend	
			CD38	HB7	Biolegend	
			lgD	IA6.2	Biolegend	

**Supplementary Table 4** List of antibodies used in T cell panel and B cell-monocyte panel for analyses of immune cell populations by flow cytometry. Antibodies were purchased from BD Biosciences and Biolegend as indicated.

Channel	Target	CLONE	Company	Code	Custom Conjugation
142Nd	SMA	1A4	Fluidigm	3141017D	No
144Nd	CD14	EPR3653	Fluidigm	3144025D	No
149Sm	CD11b	EPR1344	Fluidigm	3149028D	No
150Nd	CD15	W6D3	Fluidigm	3149026D	No
151Eu	IgD	lgD26	NovusBio	NBP2-50437	Yes
153Eu	CD24	ALB9	abcam	ab31622	Yes

154Sm	CD11c	EP1347Y	able under a CC-B abcam	4.0 International license ab216655	Yes
160Gd	CD68	KP1	Fluidigm	3159035D	No
164Dy	ARGINASE-1	D4E3M	Fluidigm	3164027D	No
165Ho	CD45RA	HI100	BioLegend	304102	Yes
166Er	CD74	LN2	Fluidigm	3166025D	No
175Lu	CD25	EPR6452	Fluidigm	3175036D	No
141Pr	CD38	EPR4106	Fluidigm	3141018D	No
143Nd	VIMENTIN	RV202	Fluidigm	3143029D	No
145Nd	CD33	POLY	Fluidigm	3145017D	No
146Nd	CD16	EPR16784	Fluidigm	3146020D	No
148Nd	Pan-Keratin	C11	Fluidigm	3148020D	No
152Sm	CD45	CD45-2B11	Fluidigm	3152016D	No
155Gd	FOXP3	236A/E7	Fluidigm	3155016D	No
156Gd	CD4	EPR6855	Fluidigm	3156033D	No
158Gd	E-CAD	24 E10	Fluidigm	3158029D	No
161Dy	CD20	H1	Fluidigm	3161029D	No
162Dy	CD8a	D8A8Y	Fluidigm	3162035D	No
168Er	CD127	EPR2955(2)	Fluidigm	3168026D	No
169Tm	COLLAGEN-I	POLY	Fluidigm	3169023D	No
170Er	CD3	POLY	Fluidigm	3170019D	No
171Yb	CD27	EPR8569	Fluidigm	3171024D	No
173Yb	CD45RO	UCHL1	Fluidigm	3173016D	No

**Supplementary Table 5** List of antibodies used in mass cytometry (CyToF) analyses for definition of immune cell subpopulations in tissue.

Pixel-level	Cell-level	It is made available undignative-BY 4.0 International license.	Populations
А	0	e-Cadherin <sup>+</sup> / Pan-Keratinin <sup>+/-</sup> / Vimentin <sup>-</sup>	Epithelium/ tumor
В	26	CD11b <sup>int</sup> / CD33 <sup>neg</sup> / CD14 <sup>int</sup> / CD15 <sup>neg</sup> / CD68 <sup>hi</sup> /CD11c <sup>int</sup> / CD74 <sup>hi</sup>	Macrophages CD74 <sup>hi</sup>
	37	CD11b <sup>int</sup> / CD33 <sup>pos</sup> / CD14 <sup>hi</sup> / CD15 <sup>neg</sup> / CD68 <sup>int</sup> /CD11c <sup>int</sup> / CD74 <sup>int</sup>	CD33+ CD14+ Monocytes
	48	CD11b <sup>int</sup> / CD33 <sup>neg</sup> / CD14 <sup>hi</sup> / CD15 <sup>neg</sup> / CD68 <sup>hi</sup> /CD11c <sup>int</sup> / CD74 <sup>int</sup>	Mono/ Macro
	60	CD11b <sup>int</sup> / CD33 <sup>pos</sup> / CD14 <sup>hi</sup> / CD15 <sup>neg</sup> / CD68 <sup>neg</sup> /CD11c <sup>neg</sup> / CD74 <sup>int</sup>	MO-MDSC
	61	CD11b <sup>int</sup> / CD33 <sup>neg</sup> / CD14 <sup>int</sup> / CD15 <sup>neg</sup> / CD68 <sup>neg</sup> /CD11c <sup>neg</sup> / CD74 <sup>int</sup>	ILC-1
	64	CD11b <sup>hi</sup> / CD33 <sup>neg</sup> / CD14 <sup>int</sup> / CD15 <sup>hi</sup> / CD68 <sup>int</sup> /CD11c <sup>neg</sup> / CD74 <sup>neg</sup>	G-MDSC
	9	CD11b <sup>int</sup> /CD33 <sup>pos</sup> /CD14 <sup>int</sup> /CD15 <sup>neg</sup> /CD68 <sup>neg</sup> /CD11c <sup>neg</sup> /CD74 <sup>neg</sup>	Monocyte
	0		Myeloid
С	0	SMA <sup>+</sup>	Fibroblast
D	33	CD3 <sup>hi</sup> / CD8 <sup>hi</sup> / CD27 <sup>neg</sup> / CD45RA <sup>neg</sup> / CD45RO <sup>int</sup> / CD127 <sup>neg</sup>	CD8 <sup>+</sup> Effector
	41	CD3 <sup>hi</sup> / CD8 <sup>hi</sup> / CD27 <sup>int</sup> / CD45RA <sup>hi</sup> / CD45RO <sup>int</sup>	CD8 <sup>+</sup> Central memory
	42	CD3 <sup>hi</sup> / CD8 <sup>hi</sup> / CD27 <sup>neg</sup> / CD45RA <sup>neg</sup> / CD45RO <sup>hi</sup>	CD8 <sup>+</sup> Effector memory
	0		CD8 <sup>+</sup> T-Cells
E	28	CD3 <sup>low</sup> / CD4 <sup>hi</sup> / CD27 <sup>neg</sup> / CD45RA <sup>neg</sup> / CD45RO <sup>int</sup> / foxP3 <sup>neg</sup> / CD25 <sup>neg</sup> / CD127 <sup>int</sup> / CD74 <sup>hi</sup>	CD4+ Effector memory
	34	CD3 <sup>hi</sup> / CD4 <sup>hi</sup> / CD27 <sup>neg</sup> / CD45RA <sup>neg</sup> / CD45RO <sup>hi</sup> / foxP3 <sup>neg</sup> / CD25 <sup>neg</sup> / CD127 <sup>int</sup> / CD74 <sup>int</sup>	CD4 <sup>+</sup> Effector memory
	49	CD3 <sup>hi</sup> / CD4 <sup>hi</sup> / CD27 <sup>neg</sup> / CD45RA <sup>neg</sup> / CD45RO <sup>int</sup> / foxP3 <sup>neg</sup> / CD25 <sup>neg</sup> / CD127 <sup>low</sup> / CD74 <sup>int</sup>	CD4 <sup>+</sup> Central memory
	50	CD3 <sup>hi</sup> / CD4 <sup>hi</sup> / CD27 <sup>int</sup> / CD45RA <sup>neg</sup> / CD45RO <sup>int</sup> / foxP3 <sup>int</sup> / CD25 <sup>low</sup> / CD127 <sup>low</sup> / CD74 <sup>int</sup>	CD4 <sup>+</sup> activated
	51	CD3 <sup>hi</sup> / CD4 <sup>hi</sup> / CD27 <sup>hi</sup> / CD45RA <sup>hi</sup> / CD45RO <sup>int</sup> / foxP3 <sup>neg</sup> / CD25 <sup>low</sup> / CD127 <sup>int</sup> / CD74 <sup>int</sup>	CD4 <sup>+</sup> Naïve
	57	CD3 <sup>hi</sup> / CD4 <sup>hi</sup> / CD27 <sup>int</sup> / CD45RA <sup>neg</sup> / CD45RO <sup>int</sup> / foxP3 <sup>hi</sup> / CD25 <sup>int</sup> / CD127 <sup>low</sup> / CD74 <sup>int</sup>	T-reg CD45RO+
	58	CD3 <sup>hi</sup> / CD4 <sup>hi</sup> / CD27 <sup>int</sup> / CD45RA <sup>int</sup> / CD45RO <sup>int</sup> / foxP3 <sup>neg</sup> / CD25 <sup>low</sup> / CD127 <sup>lint</sup> / CD74 <sup>int</sup>	CD4 <sup>+</sup> Effector
	0		CD4 <sup>+</sup> T-Cells
F	59	CD3 <sup>neg</sup> /CD20 <sup>hi</sup>	B-Cell
	0		B-Cell

**Supplementary Table 6** List of populations in imaging mass cytometry of tissue and corresponding signature. ILC = Innate Lymphoid Cells, T-reg = Regulatory T cells, SMA = Smooth Muscle Actin, MO-MDSC = Monocytic Myeloid Derived Suppressor Cells, G-MDSC = Granulocytic Myeloid Derived Suppressor Cells



medRxiv preprint doi: https://doi.org/10.1101/2021.09.22.21263942; this version posted September 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. Supplementary Figure 3 It Gratingavatrategyer forc-furtherercharacterisation of CD33+CD14+

**monocytic population using new patient cohort** CD14+CD16lowCD33+CD11b monocytes also express high levels of HLADR and CD11c.

Baseline Risk Score	Combined Risk Score
S=(-0.278426819992)*MDSC.of.CD33c1	S=(1.615133865474)*Site.Hypopharynx
+ (-0.060298841476)*B.cell.DN.c1	+ (-0.224319089905)*MDSC.of.CD33c1
+ (0.053313459288)*Age	+ (-0.057227235933)*CD4.Memory.Tregs.c1
- (1.142256112742)	+ (0.561268824195)*T.cell.CD3.lfc
	+ (1.290959733907)*CD8.Cent.Mem.lfc
	+ (-1.498015339199)*CD8.Effect.Mem.lfc
	- (-3.790127113280)

**Supplementary Figure 4** Risk score signatures for use with raw covariate values, with missing data imputed with the study mean from Supplementary Table 1. Use these equations to calculate the risk score for each patient.