

eLife's Review Process

eLife works to improve the process of peer review so that it more effectively conveys the assessment of expert reviewers to authors, readers and other interested parties. In the future we envision a system in which research is first published as a preprint and the outputs of peer review are the primary way research is assessed, rather than journal title.

Our editorial process produces two outputs: i) an assessment by peers designed to be posted alongside a preprint for the benefit of the readers; ii) detailed feedback on the manuscript for the authors, including requests for revisions and suggestions for improvement.

Therefore we want to change how we construct and write peer reviews to make them useful to both authors and readers in a way that better reflects the work you put into reading and thinking about a paper.

eLife reviews now have three parts:

- An **evaluation summary** (in two or three sentences) that captures the major conclusions of the review in a concise manner, accessible to a wide audience.
- A **public review** that details the strengths and weaknesses of the manuscript before you, and discusses whether the authors' claims and conclusions are justified by their data.
- A set of private **recommendations for the authors** that outline how you think the science and its presentation could be strengthened.

All three sections will be used as the basis for an eLife publishing decision, which will, as always, be made after a consultation among the reviewers and editor. Each of the **public reviews** will be published (anonymously) alongside the preprint, together with a response from the authors if they choose. In the case of papers we reject after review, the authors can choose to delay posting until their paper has been published elsewhere.

If this is your first time going through this new process, we ask that you take some time to read our [Reviewer Guide](#), which discusses how we see each section will be used, what it should contain, and what we hope it accomplishes. And we remind you that, with the shift of reviews from private correspondence to public discourse, it is more important than ever that reviews are written in a **clear and constructive manner** appropriate for a public audience and mindful of the impact language choices might have on the authors.

Information about the manuscript

A Machine Learning Approach Yields An Immune-Based Multiparameter Predictive Signature to Systemic Therapy in Advanced Head and Neck Cancer

Tracking no: 23-08-2021-RA-eLife-73288

Competing interests: Paul Barber: PRB is a shareholder of Nano Clinical Ltd Kenrick Ng: KN has received honoraria from Pfizer, GSK/Tesaro and Boehringer Ingelheim, and has had travel/accommodation/expenses paid for by Tesaro. Shahram Kordasti: SK has received research funding in the form of a grant from Novartis and Celgene Jana Doyle: JD is in employment with Daichi Sankyo, and has stock and other ownership interests, research funding within Daichi Sankyo and has had travel/accommodation/expenses paid for by Daichi Sankyo Jon Greenberg: JG is in employment with Daichi Sankyo, and has stock and other ownership interests, research funding within Daichi Sankyo and has had travel/accommodation/expenses paid for by Daichi Sankyo Kevin Harrington: KH has received honoraria from Amgen; Arch Oncology; AstraZeneca; Boehringer-Ingelheim; Bristol-Myers Squibb; Codiak; Inzen; Merck; MSD; Pfizer; Replimune and is on a speakers' bureau for Amgen, AstraZeneca; Bristol-Myers Squibb; Merck, MSD; Pfizer. KH has also received research funding from AstraZeneca, Boehringer-Ingelheim, MSD and Replimune. Martin Forster: MDF has received institutional 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, Guardant Health, Merck, MSD, Nanobiotix, Novartis, Oxford VacMedix, Pfizer, Roche, Takeda, UltraHuman Anthony Coolen: ACCC has stock and other ownership interests with Saddle Point Science Limited. Tony Ng: TN has received research funding from Astrazeneca and Daichi Sankyo. TN is a founder and shareholder in Nano Clinical Ltd, and PRB is a shareholder.

Author contributions:

Paul Barber: Conceptualization; Data curation; Software; Formal analysis; Investigation; Methodology Fabian Flores-Borja: Data curation; Formal analysis; Investigation; Methodology Giovanna Alfano: Data curation; Formal analysis; Investigation Kenrick Ng: Formal analysis; Methodology; Writing - original draft; Writing - review and editing Gregory Weitsman: Formal analysis; Investigation Luigi Docketti: Formal analysis; Investigation Rami Mustapha: Conceptualization; Formal analysis; Methodology Felix Wong: Formal analysis; Investigation Jose Vicencio: Formal analysis; Investigation; Methodology Myria Galazi: Data curation; Investigation James Opzoomer: Data curation; Formal analysis; Investigation James

Arnold: Supervision; Writing - review and editing Shahram Kordasti: Formal analysis; Supervision Jana Doyle: Resources; Supervision; Funding acquisition; Project administration Jon Greenberg: Resources; Funding acquisition; Project administration; Writing - review and editing Magnus Dillon: Methodology; Project administration Kevin Harrington: Supervision; Methodology; Project administration; Writing - review and editing Martin Forster: Conceptualization; Supervision; Methodology; Writing - review and editing Anthony Coolen: Conceptualization; Data curation; Formal analysis; Supervision; Methodology Tony Ng: Conceptualization; Resources; Data curation; Supervision; Funding acquisition; Writing - review and editing

Data Availability:

The data generated in this study are available upon request from the corresponding author.

N/A

Ethics:

Human Subjects: Yes Ethics Statement: 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: 15/LO/1670). Approval to procure and process a separate cohort of blood samples from patients at risk of developing lung cancer was also obtained (IRAS ID: 261766). Clinical Trial: NCT02633800 Animal Subjects: No

1 **A Machine Learning Approach Yields An Immune-Based**
2 **Multiparameter Predictive Signature to Systemic Therapy in Advanced**
3 **Head and Neck Cancer**

4
5 Paul R Barber^{†1,2}, Fabian Flores-Borja^{‡*3}, Giovanna Alfano^{‡4}, Kenrick Ng^{†1}, Gregory Weitsman⁴,
6 Luigi Dolcetti⁴, Rami Mustapha⁴, Felix Wong⁴, Jose M Vicencio^{1,4}, Myria Galazi¹, James W
7 Opzomer⁵, James N Arnold⁵, Shahram Kordasti⁶, Jana Doyle⁷, Jon Greenberg⁷, Magnus
8 T.Dillon⁸, Kevin J. Harrington⁸, Martin D. Forster¹, Anthony C C Coolen^{9,10}, and Tony Ng^{1,3,4†}.

- 9
10 1. UCL Cancer Institute, Paul O’Gorman Building, University College London, London, UK
11 2. Comprehensive Cancer Centre, School of Cancer & Pharmaceutical Sciences, King’s
12 College London, London, UK.
13 3. Breast Cancer Now Research Unit, School of Cancer & Pharmaceutical Sciences, King’s
14 College London, London, UK.
15 4. Richard Dimpleby Laboratory of Cancer Research, School of Cancer & Pharmaceutical
16 Sciences, King’s College London, London, UK.
17 5. Tumor Immunology Group, School of Cancer & Pharmaceutical Sciences, King’s College
18 London, London, UK
19 6. Systems Cancer Immunology, School of Cancer & Pharmaceutical Sciences, King’s College
20 London, London, UK
21 7. Daichii Sankyo Incorporated, New Jersey, USA
22 8. The Institute of Cancer Research, London, UK
23 9. Institute for Mathematical and Molecular Biomedicine, King’s College London, UK
24 10. Saddle Point Science Ltd, London, UK
25

26 ‡ These authors contributed equally to this work.

27 * Current address: Centre for Immunobiology and Regenerative Medicine, Barts & The
28 London School of Medicine and Dentistry, Queen Mary University of London, London, UK

29 † To whom correspondence should be addressed:

30 **Prof. Tony Ng**

31 Tel no: +44 (0) 20 7848 8056 (KCL)

32 E-mail: tony.ng@kcl.ac.uk, t.ng@ucl.ac.uk

33

34 **Word Count:** 4833 words (excluding figure legends and references)

35 **Abstract:** 151 words

36 **Running Title:** Immune-Based Predictive Signature for Head and Neck Cancer

37

38

39

40

41

42

43

44

45

46

47

48

49

50 **List of abbreviations:**

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

70 **ABSTRACT**

71 Advanced Head and Neck Squamous cell Cancer (HNSCC) is associated with a poor
72 prognosis. The biological framework to predict which patients would respond to treatment is not
73 well established. Using blood samples from HNSCC patients within a Phase 2 clinical trial, we
74 obtained a broad range of clinical and laboratory-based information at baseline and after they
75 received one cycle of chemotherapy-based treatment. By using a Bayesian-based machine
76 learning modelling approach, we identified a set of six predominantly immune-based markers
77 which predict benefit to treatment. We also demonstrated that longitudinal sampling of
78 peripheral blood generates a stronger predictive signature than biomarkers derived prior to
79 treatment. Further, we identify and characterize a novel monocytic population, termed
80 CD33+CD14+ monocytes, which may have implications for response to treatment in this patient
81 population. The data presented identifies a novel means of longitudinal blood sampling to
82 predict response early during the treatment course of a HNSCC patient.

83

84

85

86

87

88

89

90

91

92

93

94 INTRODUCTION

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

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

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

116 Efforts to develop a machine learning model to stratify survival risk by combining genetic
117 and clinicopathological characteristics have revealed some success in advanced oral squamous

118 cell carcinoma(7). We hypothesize that a multimodal analysis, taking into account both
119 clinicopathological and laboratory-based biological covariates at different timepoints, would
120 provide better predictive value.

121 We prospectively collected peripheral blood samples from a Phase 2 trial in R/M HNSCC
122 (NCT02633800)(8), which utilized a modified EXTREME regimen as a backbone, and
123 conducted a parallel exploratory analysis with the aim of generating a biomarker signature
124 which would predict outcomes to treatment. We hypothesized that the detailed definition of a
125 broad immune cell signature could contribute to the development of assays employing liquid
126 biopsies to predict clinical outcomes. We also incorporated the analysis of two circulating
127 microRNAs (miRNAs); miR-21-5p and miR-142-3p, which have previously demonstrated
128 prognostic and predictive utility(9, 10). As the trial investigated the efficacy of an anti-ErbB3
129 antibody, patritumab, administered alongside an anti-EGFR antibody, we simultaneously
130 analyzed EGFR-ErbB3 dimerization using Förster Resonance Energy Transfer (FRET) and
131 included it in our analysis.

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

137

138 **MATERIALS AND METHODS**

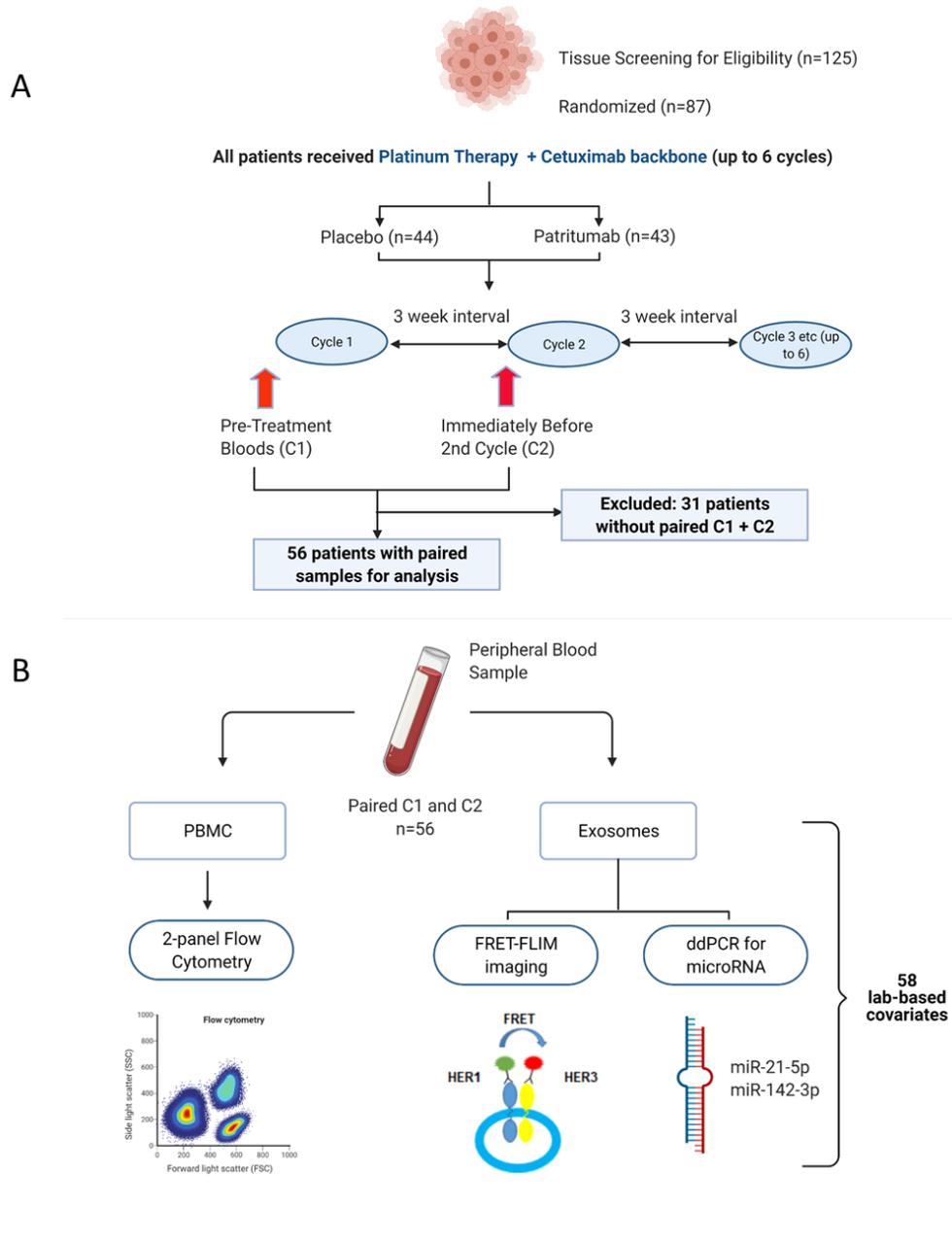
139 **Study Design**

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

149 PBMC samples were analyzed using Flow Cytometry to generate unique immunological
150 subpopulations. Exosomes were extracted from the serum and analyzed for EGFR-ErbB3
151 dimerization and miRNA-21-5p and miRNA-142-3p (Figure 1B). These analyses yielded a total
152 of 29 unique biological covariates. Each covariate was obtained in pairs (C1 and C2),
153 generating a total of 58 laboratory-based covariates for the multivariate analysis (Figure 1B). To
154 mitigate individual baseline variations between patients, we evaluated changes between C1 and
155 C2 (in the form of log-fold change, LFC, of the variable of interest) instead of absolute values of
156 those parameters. A list of the laboratory-based and clinical covariates is provided in
157 Supplementary Table 2.

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

164 Written informed consent was obtained. Approval was obtained from ethics committees
 165 (Research Ethics Committee reference: 15/LO/1670).
 166



167
 168

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

173

174 **Statistical Analysis**

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

184 The relative efficiency of the predictive model was assessed by using C-index (a metric
185 proposed by Harrell(12) to evaluate the accuracy of predictions made by an algorithm) and rank
186 correlation of the signature-generated risk scores with survival time. The number of significant
187 covariates in each prediction signature was determined with the aim of avoiding overfitting of the
188 signature to the study data using the “batch regression” option of the Saddle Point Signature
189 software (Saddle Point Science Ltd., London, UK), according to methods that were previously
190 published(13, 14). This is particularly important with small number of patients where an
191 independent test set is not possible. Systematic iterative covariate rejection and cross-validation
192 (5000 iterations) allowed for the selection of an optimal covariate set to avoid overfitting though
193 inclusion of too many covariates. The optimal set can be chosen in two ways, either based on

194 the peak prediction performance of cross-validation, or the more stringent method that equally
195 penalizes validation performance and overfitting (defined as the deviation between training and
196 validation performance). All signatures presented were chosen using the more stringent criterion
197 and data for all covariates is also presented for the purposes of identifying covariates that may
198 be important but do not quite meet the criterion. The regression included covariates
199 representing the missingness of data to account for the possibility that patient or sample
200 selection/rejection (for any reason) is biased with respect to outcome and therefore could be
201 informative. The importance and significance of covariates can be judged by their assigned beta
202 value (β) in the proportional hazards model, and corresponding hazard ratio (HR) equal to $e^{2\beta}$. A
203 negative beta value reflects a lower risk of developing an event. The Signature software
204 additionally judges the performance of similar randomized data, which most often has beta
205 values around zero and within a critical range, such that any real covariate that has a beta value
206 outside this critical range can be judged to be performing significantly better than randomized
207 data. This adds additional confidence in the absence of an independent validation test set.

208 **Flow Cytometry**

209 Frozen PBMC samples were thawed and stained with Fixable viability dye (Yellow
210 Live/Dead™, Fisher Scientific) followed by two different panels of membrane markers. A panel
211 for T cells included CD3, CD4, CD8, CD25, CD45RO, CD127, CCR7, and HLA-DR. A panel for
212 B cells and monocytes included CD3, CD19, CD24, CD38, CD27, IgD, CD33, CD11b, CD14
213 and CD16 (full list of both antibody panels in Supplementary Table 4). These two panels allow
214 definition of immune cell populations as described in Supplementary Figure 2. Patients' samples
215 and corresponding Fluorescence Minus One (FMO) Controls were acquired in a Fortessa II flow
216 cytometer (BD, Berkshire, UK) and analyzed with FlowJo software (Tree Star).

217 **Isolation of Serum exosomes**

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

225 **RNA Extraction and miRNA Expression Analysis**

226 RNA from cancer patients' serum exosomes was extracted using the TRIzol™ Plus
227 RNA Purification Kit (Thermo Fisher, UK) according to the manufacturer' instructions.
228 Quantification of gene expression in circulating exosomes was performed by ddPCR (Bio-Rad
229 QX100 system). Normalization of the RNA, between cycle 1 and cycle 2 therapy of each patient,
230 was performed using the expression levels of the housekeeping gene 18S (Assay ID,
231 Hs99999901_s1). For each sample, equal volume of RNA was used as template and cDNA
232 synthesis performed using the SuperScript® VILO™ MasterMix (Thermo Fisher, UK) according
233 to the manufacturer' instructions. MicroRNAs were reverse-transcribed individually using the
234 TaqMan™ MicroRNA Reverse Transcription Kit (Thermo Fisher, UK). For each sample,
235 the normalized amount of RNA was reverse transcribed in a 15 µl reaction using the standard
236 protocol and primers specific for each miRNA: miR-21-5p (assay ID, 000397), miR-142-3p
237 (assay ID, 000464). Then 7.5 µl of cDNA was added to a 20 µl reaction containing 12.5 µl 2X
238 ddPCR Supermix for Probes (Bio-Rad) and 1 µl 20X TaqMan miRNA PCR primer probe set;
239 each reaction was carried out in duplicate. Thermo cycling conditions were as it follows: 95 °C

240 for 10 min, then 50 cycles of 95 °C for 10 sec and 61 °C for 30 sec and a final inactivation step
241 at 98 °C for 12 min. PCR products were analyzed using the QuantaSoft™ Software (Bio-Rad).

242 **ErbB3-EGFR Dimer Quantification in Exosomes**

243 Exosomes were imaged on an 'Open' Fluorescence Lifetime Imaging Microscopy (FLIM)
244 system (16). Analysis was performed with the TRI2 software (v2.7.8.9, CRUK/MRC Oxford
245 Institute for Radiation Oncology, Oxford) as described previously (17, 18). Interfering effects of
246 autofluorescence were minimized with a lifetime filtering algorithm and the FRET efficiency
247 value for each patient calculated by: $FRET=1-\frac{\tau_{DA}}{\tau_D}$, where t_D and t_{DA} are the average lifetime of
248 Alexa Fluor546 in the matching donor (D) and donor-acceptor (DA) images.

249

250 **Imaging Mass Cytometry**

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

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

257 Ablation and data acquisition were performed on a Fluidigm Hyperion located within our
258 Biomedical Research Centre. Imaging analysis was performed using the following R packages:
259 RandomForest for classification and regression, Raster and SF for image manipulation and
260 segmentation. Scripts are available upon request from the corresponding author.

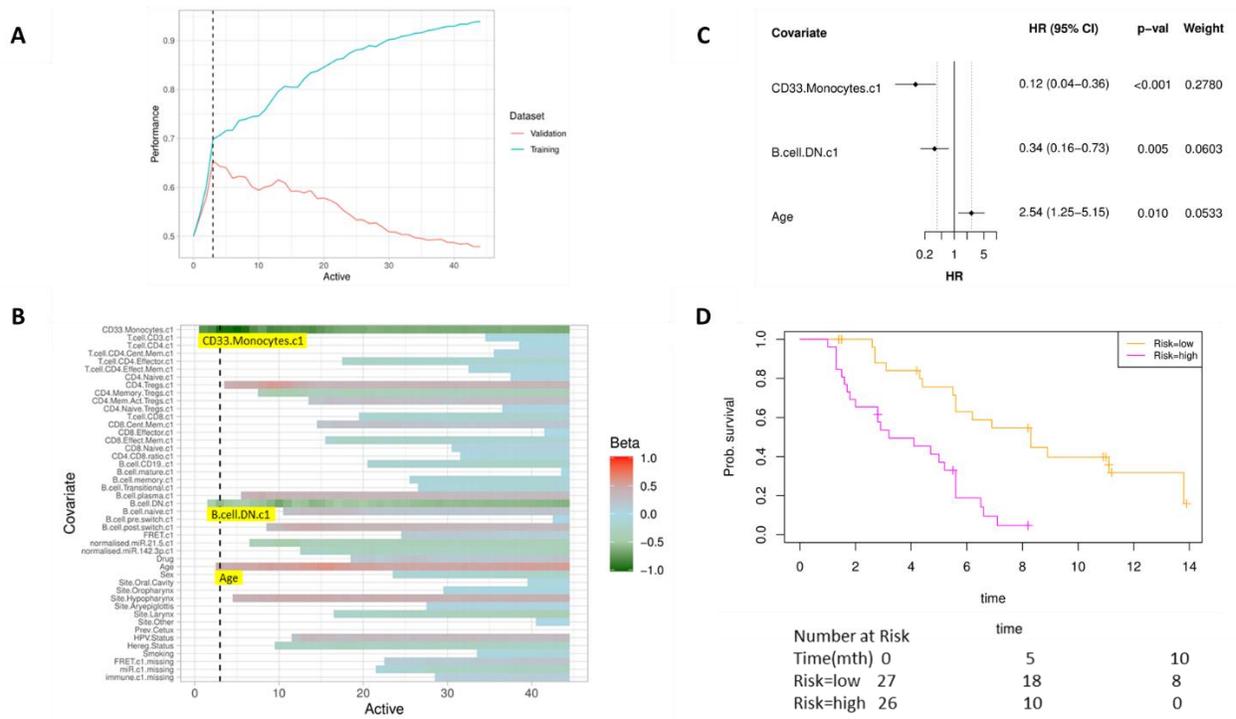
261 **RESULTS**

262 **The Model with Baseline Covariates Reveal Immune Subpopulations and Age Predict**
263 **PFS**

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

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

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



283

284 **Figure 2 High baseline CD33+CD14+ monocytes and Double Negative B cells Predict Overall Survival (A)**
 285 Covariates were ranked for importance and selected by a proportional hazards regression model with cross-
 286 validation. (B) Proportional hazards regression revealed three covariates which exceeded the beta critical value –
 287 CD33+CD14+ monocytes, Double Negative B Cells, and Age. (C) Forest plot of the three covariates within
 288 Progression Free Survival risk score with dotted line indicating the range, around 1, of typical random covariates. (D)
 289 Progression Free Survival Risk Signature Performance, low risk score (n=27) and high risk score (n=26). Log rank P-
 290 value = 6.0 e-5, with numbers at risk demonstrated under Kaplan-Meier curve. The multivariate analysis resulted in
 291 risk signatures that are linear combinations of weighted covariates. Their ability to predict outcome is demonstrated
 292 with data split by signature value.

293

294

295

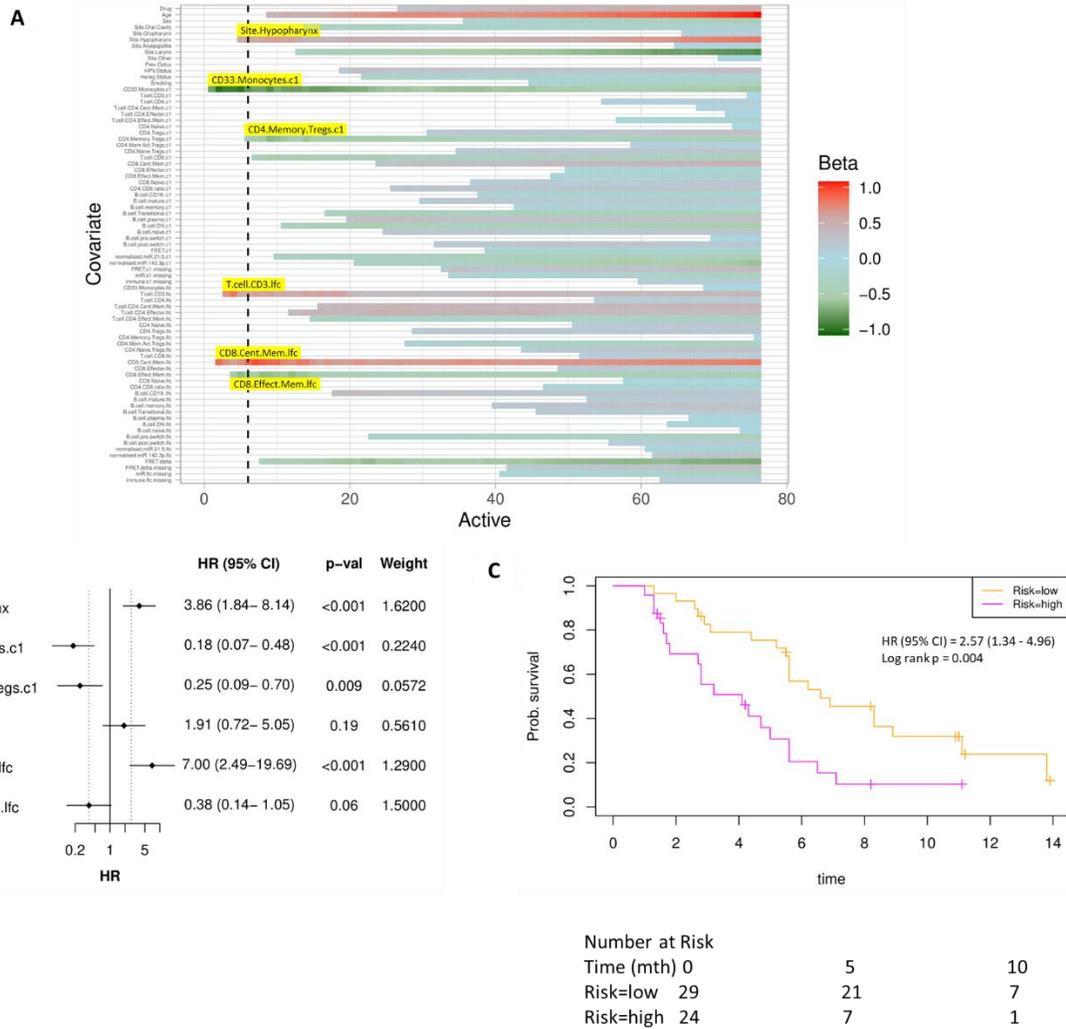
296

297 **Incorporating laboratory-based covariates after one cycle of treatment improves ability to**
298 **predict PFS benefit**

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

303 As before, we used a proportional hazards regression to determine a set of variables which
304 predict PFS. A total of six covariates were identified – three immune subpopulations with
305 negative beta values and hence are associated with better survival, ie baseline CD33+CD14+
306 monocytes, baseline CD4 memory regulatory T cells (HLA-DR⁻CD45RO⁺Tregs) and an increase
307 in CD8 effector memory T cells (CD45RO⁺CCR7⁻). An increase in two subpopulations, CD8
308 Central Memory T Cells (CD45RO⁺CCR7⁺) and CD3 T cells, was associated with inferior PFS.
309 The hypopharyngeal primary tumor site was also associated with a poorer PFS (Figure 3A and
310 B).

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



319

320 **Figure 3 Model Incorporating Laboratory Changes After One Cycle of Treatment Exhibit Improved Predictive Value (A)**

321 Proportional hazards regression revealed five immune subpopulations which exceeded the beta critical value – Baseline CD33+CD14+

322 monocytes, Baseline CD4 Memory Regulatory T Cells, LFC of CD8 effector memory T cells, LFC of CD8 central memory T cells

323 and LFC of CD3 T cells. The primary tumor site of hypopharynx also featured in the signature. A negative beta value is associated

324 with lower risk score and hence better progression free survival. LFC = Log Fold Change (B) Forest plot of the three covariates

325 within Progression Free Survival risk score (C) Progression Free Survival Risk Signature Performance, low risk score (n=29) and

326 high risk score (n=24). Log rank P-value = 0.004, with numbers at risk demonstrated under Kaplan-Meier curve.

327

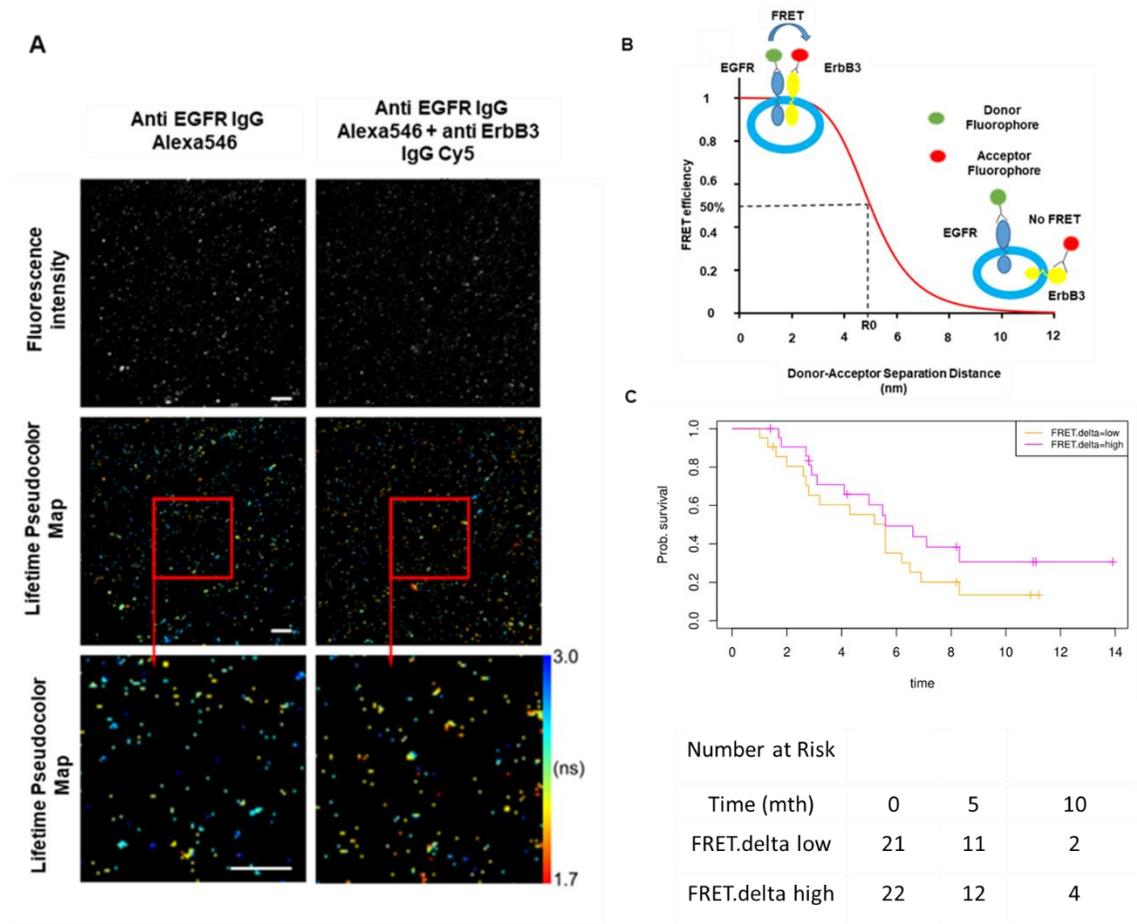
328

329 **EGFR-ErbB3 FRET may contribute to predictive signature**

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

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

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



346

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

352

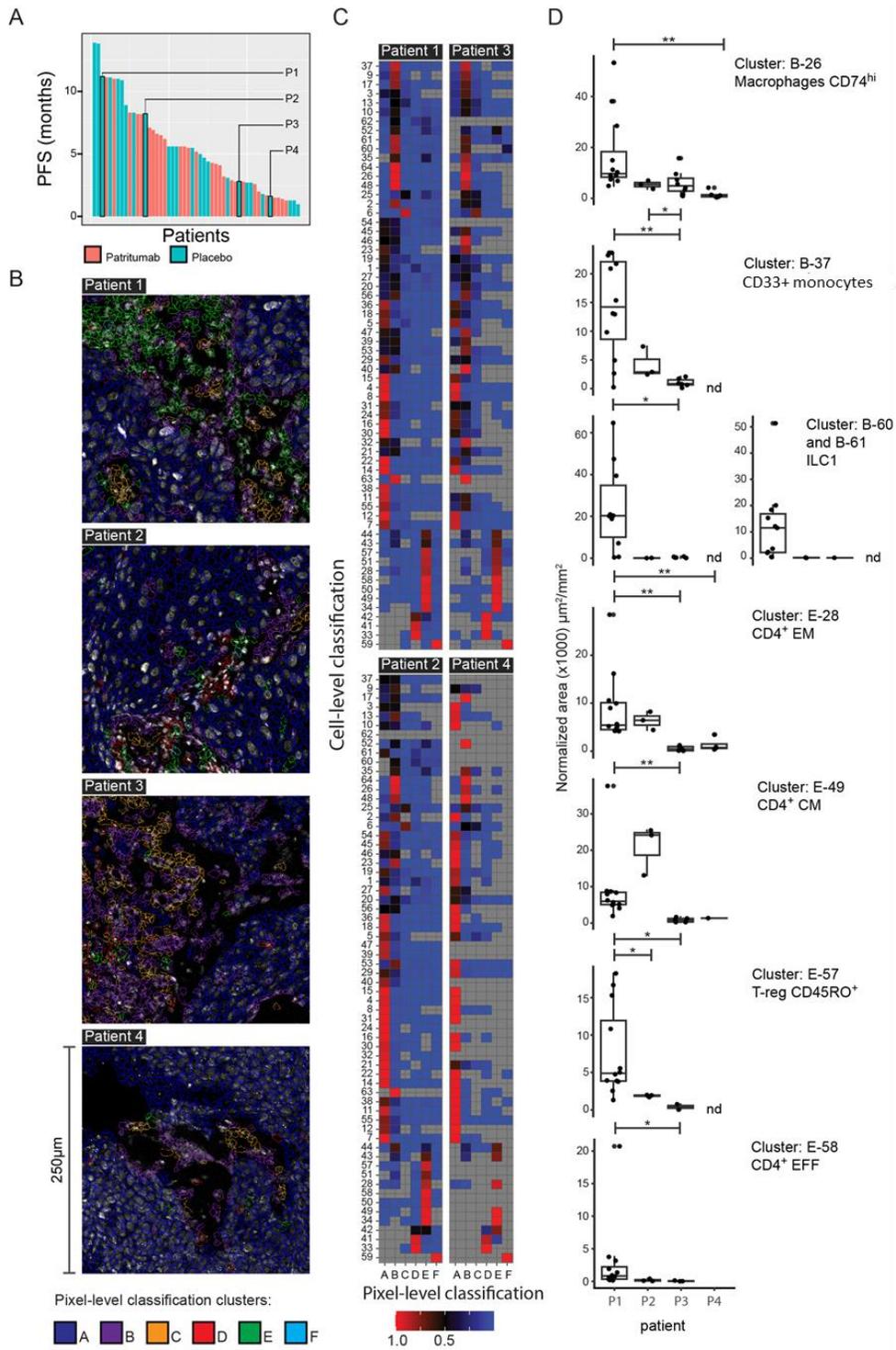
353 **Imaging Mass Cytometry of Tissue Reveals Correlation of CD33+CD14+ Myeloid Cell**
 354 **Subpopulation Between Tissue and Peripheral Blood**

355 Having established that immune subsets in peripheral blood predict therapeutic response within
 356 a multivariate signature, we subsequently explored the relationship between the immune

357 findings in peripheral blood with tumor infiltrating leukocytes (TILs). We obtained sufficient
358 tissue from the biopsy at trial enrolment for in-depth profiling by imaging mass cytometry from
359 four patients.

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

368 A total of 7 cell clusters identified on tissue mass cytometry were significantly different between
369 the four patients. CD33+CD14+ monocytes (cluster B-37) which featured in both baseline and
370 combined signatures, exhibited a diminishing trend of abundance across patients 1 to 4. Having
371 observed this trend, we subsequently correlated the levels of this population in tissue with
372 blood. The proportion of CD33+CD14+ monocytes in the blood was 17.4%, 5.39%, 2.0% and
373 2.47% for patients 1-4 respectively, suggesting a meaningful concordance between the levels of
374 this subpopulation in the patient tissue and peripheral blood.



375

376

377

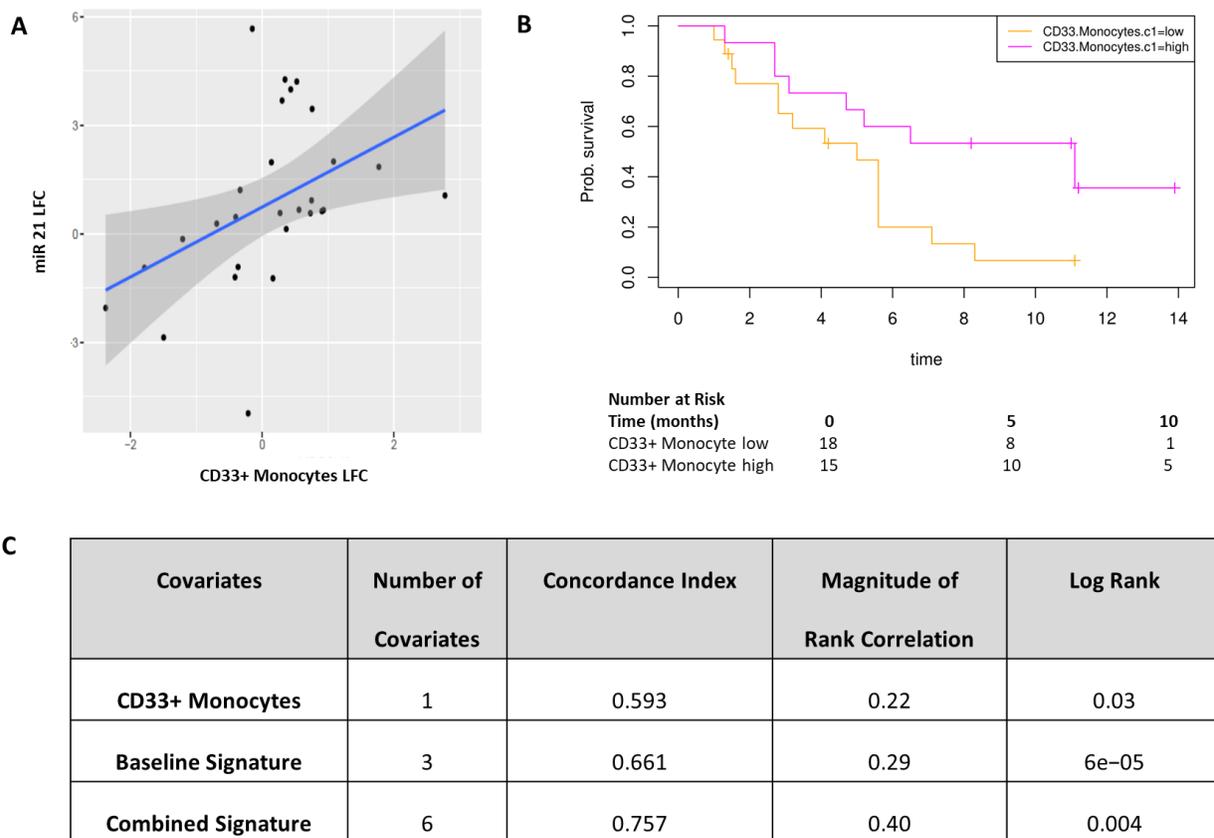
378 **Figure 5 Image Mass Cytometry of Four Patient Samples Using a 29 Marker Panel Analysis** (A) Waterfall plot of
379 Progression Free Survival showing the four patients used for imaging mass cytometry analysis (B) Representative
380 image of the nuclear staining overlaid with pixel-level classification (C) Heatmaps representing the distribution of cell
381 phenotypes for each patient, as expressed by a two-phase clustering conducted at pixel level (columns) and at cell
382 level (rows). Data represent row-normalized areas. Red tiles, which represent hot spots of classification concordance,
383 are further described in panel D (D) Differential analysis of the classification hotspots presented in panel C,
384 highlighting cell populations which were significantly different between patients. Data are presented as sum of areas
385 of positive cells normalized to the total area of the ROI and expressed as μm^2 ($\times 1000$) per mm^2 . Statistics: *adj-
386 $p < 0.05$, **adj- $p < 0.01$, ****adj- $p < 0.001$, Pairwise Wilcoxon Rank Sum Tests with Benjamini correction, $n \geq 3$.

387 **CD33+CD14+ Monocyte Population have high HLA-DR expression and univariate**
388 **predictive value**

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

397 MiRNA signatures have been implicated as a useful classifier for myeloid cell subsets(20). By
398 correlating the miRNA changes in our study with this monocytic subpopulation, a significant
399 correlation was identified between the log fold changes of miR-21-5p with the corresponding log
400 fold changes of CD33+CD14+ monocytes (Pearson's $r=0.4343$, $p=0.02092$, Figure 6A). We also
401 investigated the potential of baseline CD33+CD14+ monocyte levels at predicting PFS. Figure
402 6B illustrates a Kaplan-Meier curve of PFS by median CD33+CD14+ monocyte level, generating

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



408
 409 **Figure 6 CD33+CD14+ Monocytes have some predictive value but predictive ability of signature is maximized with**
 410 **longitudinal sampling of peripheral blood** (A) (B) Correlation between miRNA 21 fold change and CD33+ monocyte fold change
 411 after one cycle of treatment. Correlation coefficient of 0.4343, p value 0.02092. miR21: microRNA-21-5p, lfc: log-fold change (B)
 412 Kaplan-Meier Curve of PFS split by median CD33+CD14+ value, log rank p-value 0.03 (C) Table summarises C-index, rank
 413 correlation and log-rank p value based on type and number of covariates. There is an increasing C-index and rank correlation upon
 414 the strict selection of additional covariates into the predictive signature, with the strongest signature incorporating six covariates
 415 combining values from baseline and after one cycle of treatment.

416 **DISCUSSION**

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

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

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

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

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

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

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

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

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

482 **FUNDING**

483 This work was supported by a grant from Daiichi Sankyo Inc ('Identification of Non-
484 Invasive Treatment Stratification and Longitudinal Monitoring Markers for Patritumab/Cetuximab
485 Combination Therapy'). This work was also supported by Cancer Research UK funding support
486 to King's College London – UCL Comprehensive Cancer Imaging Centre (CR-UK & EPSRC),
487 Cancer Research UK King's Health Partners Centre at King's College London, and Cancer
488 Research UK UCL Centre; University College London (PRB) – Early Detection Award

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

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

497 MDF is supported by the UCL/UCLH NIHR Biomedical Research
498 Centre and runs early phase studies in the NIHR UCLH Clinical Research
499 Facility supported by the UCL ECMC. MTD and KH acknowledge funding support from The
500 Institute of Cancer Research/Royal Marsden Hospital NIHR Biomedical Research Centre.

501 **DATA AVAILABILITY STATEMENT**

502 The data generated in this study are available upon request from the corresponding author.

503 **NOTES**

504 **Affiliations of authors**

505 Breast Cancer Now Research Unit, School of Cancer & Pharmaceutical Sciences, King's
506 College London, London, UK (FF-B,TN); Richard Dimpleby Laboratory of Cancer Research,
507 School of Cancer & Pharmaceutical Sciences, King's College London, London, UK
508 (GA,GW,FW,RM,TN); Comprehensive Cancer Centre, School of Cancer & Pharmaceutical
509 Sciences, King's College London, London, UK (PRB); UCL Cancer Institute, Paul O'Gorman
510 Building, University College London, London, UK (PRB,JMV,MG,KN,JB,MDF,TN); Tumor

511 Immunology Group, School of Cancer & Pharmaceutical Sciences, King's College London,
512 London, UK (JWO, JNA), Systems Cancer Immunology, School of Cancer & Pharmaceutical
513 Sciences, King's College London, London, UK (SK), Daiichi Sankyo Incorporated, New Jersey,
514 USA (JD, JG); The Institute of Cancer Research/The Royal Marsden Hospital National Institute
515 for Health Research Biomedical Research Centre, London, UK (KH, MTD); Institute for
516 Mathematical and Molecular Biomedicine, King's College London, UK (ACC), Saddle Point
517 Science Ltd, London, UK (ACC), Centre for Immunobiology and Regenerative Medicine, Barts &
518 The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
519 (Current Affiliation for FF-B) ; Barts Health NHS Trust, London, UK (Current Affiliation for MG).

520 **Conflict of Interests**

521 KN has received honoraria from Pfizer, GSK/Tesaro and Boehringer Ingelheim, and has
522 had travel/accommodation/expenses paid for by Tesaro. MDF has received institutional
523 research funding from AstraZeneca, Boehringer-Ingelheim, Merck and MSD and serves in a
524 consulting or advisory role to Achilles, Astrazeneca, Bayer, Bristol-Myers Squibb, Celgene,
525 Guardant Health, Merck, MSD, Nanobiotix, Novartis, Oxford VacMedix, Pfizer, Roche, Takeda,
526 UltraHuman. KH has received honoraria from Amgen; Arch Oncology; AstraZeneca;
527 Boehringer-Ingelheim; Bristol-Myers Squibb; Codiak; Inzen; Merck; MSD; Pfizer; Replimune and
528 is on a speakers' bureau for Amgen, AstraZeneca; Bristol-Myers Squibb; Merck, MSD; Pfizer.
529 KH has also received research funding from AstraZeneca, Boehringer-Ingelheim, MSD and
530 Replimune.

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

536 FW was initially funded by Daichii Sankyo as research assistant to conduct laboratory
537 work in the context of the translational aspect of this trial. SK has received research funding in
538 the form of a grant from Novartis and Celgene. TN has received research funding from
539 Astrazeneca and Daichii Saky. TN is a founder and shareholder in Nano Clinical Ltd, and PRB
540 is a shareholder.

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

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

544 **ETHICS AND PERMISSIONS**

545 Written informed consent was obtained for all patients who participated in the Phase 2 clinical
546 trial. Approval was obtained from ethics committees (Research Ethics Committee reference:
547 15/LO/1670). Approval to procure and process a separate cohort of blood samples from patients
548 at risk of developing lung cancer was also obtained (IRAS ID: 261766).

549 **ACKNOWLEDGEMENTS**

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

553 **PRIOR PRESENTATIONS**

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

556 **LIST OF SUPPLEMENTARY MATERIALS**

557 Supplementary Figure 1: Kaplan Meier Curve of Progression Free Survival in Study Cohort
558 Supplementary Table 1: Demographic and Laboratory-Based Values of Patients Within Study
559 Supplementary Table 2: List of covariates entered into Bayesian Multivariate Analysis
560
561 Supplementary Table 3: Demographic and Laboratory-Based Values of Patients Separated by
562 Arm of Treatment Within Trial
563 Supplementary Figure 2: Gating Strategies For Definition of Peripheral Blood Immune
564 Populations
565 Supplementary Table 4: List of antibodies used in T cell panel and B cell-Monocyte Panel for
566 Flow Cytometry
567 Supplementary Table 5: List of Antibodies Used in Imaging Mass Cytometry (CyTOF) Analysis
568 for Definition of Immune Subpopulations
569 Supplementary Table 6: List of populations in imaging mass cytometry of tissue and
570 corresponding signature
571 Supplementary Figure 3: Gating strategy for further characterization of CD33+CD14+ monocytic
572 population using new patient cohort

573 REFERENCES

- 574 1. Burtneß B, Harrington KJ, Greil R, Soulieres D, Tahara M, de Castro G, Jr., et al. Pembrolizumab alone or
575 with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the
576 head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. *Lancet*. 2019;394(10212):1915-28.
- 577 2. Van Cutsem E, Kohne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A, et al. Cetuximab and
578 chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med*. 2009;360(14):1408-17.
- 579 3. Bossi P, Bergamini C, Siano M, Cossu Rocca M, Sponghini AP, Favales F, et al. Functional Genomics
580 Uncover the Biology behind the Responsiveness of Head and Neck Squamous Cell Cancer Patients to Cetuximab.
581 *Clin Cancer Res*. 2016;22(15):3961-70.
- 582 4. You GR, Cheng AJ, Lee LY, Huang YC, Liu H, Chen YJ, et al. Prognostic signature associated with
583 radioresistance in head and neck cancer via transcriptomic and bioinformatic analyses. *BMC Cancer*. 2019;19(1):64.
- 584 5. Cheerla A, Gevaert O. Deep learning with multimodal representation for pancancer prognosis prediction.
585 *Bioinformatics*. 2019;35(14):i446-i54.
- 586 6. Huang Z, Zhan X, Xiang S, Johnson TS, Helm B, Yu CY, et al. SALMON: Survival Analysis Learning
587 With Multi-Omics Neural Networks on Breast Cancer. *Front Genet*. 2019;10:166.

588 7. Tseng YJ, Wang HY, Lin TW, Lu JJ, Hsieh CH, Liao CT. Development of a Machine Learning Model for
589 Survival Risk Stratification of Patients With Advanced Oral Cancer. *JAMA Netw Open*. 2020;3(8):e2011768.

590 8. Forster MD, Dillon MT, Kocsis J, Remenar E, Pajkos G, Rolland F, et al. Patritumab or placebo, with
591 cetuximab plus platinum therapy in recurrent or metastatic squamous cell carcinoma of the head and neck: A
592 randomised phase II study. *Eur J Cancer*. 2019;123:36-47.

593 9. Summerer I, Unger K, Braselmann H, Schuettrumpf L, Maihoefer C, Baumeister P, et al. Circulating
594 microRNAs as prognostic therapy biomarkers in head and neck cancer patients. *Br J Cancer*. 2015;113(1):76-82.

595 10. Vahabi M BG, Di Agostino, S. MicroRNAs in head and neck squamous cell carcinoma: a possible
596 challenge as biomarkers, determinants for the choice of therapy and targets for personalized molecular therapies.
597 *Transl Cancer Res*. 2021;10(6):3090-110.

598 11. Grigoriadis A, Gazinska P, Pai T, Irhsad S, Wu Y, Millis R, et al. Histological scoring of immune and
599 stromal features in breast and axillary lymph nodes is prognostic for distant metastasis in lymph node-positive breast
600 cancers. *J Pathol Clin Res*. 2018;4(1):39-54.

601 12. Harrell FE, Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved
602 prognostic prediction. *Stat Med*. 1984;3(2):143-52.

603 13. Barber PR, Weitsman G, Lawler K, Barrett JE, Rowley M, Rodriguez-Justo M, et al. HER2-HER3
604 Heterodimer Quantification by FRET-FLIM and Patient Subclass Analysis of the COIN Colorectal Trial. *J Natl*
605 *Cancer Inst*. 2020;112(9):944-54.

606 14. Shalabi A, Inoue M, Watkins J, De Rinaldis E, Coolen AC. Bayesian clinical classification from high-
607 dimensional data: Signatures versus variability. *Stat Methods Med Res*. 2018;27(2):336-51.

608 15. Monypenny J, Milewicz H, Flores-Borja F, Weitsman G, Cheung A, Chowdhury R, et al. ALIX Regulates
609 Tumor-Mediated Immunosuppression by Controlling EGFR Activity and PD-L1 Presentation. *Cell Rep*.
610 2018;24(3):630-41.

611 16. Barber PR, Tullis ID, Pierce GP, Newman RG, Prentice J, Rowley MI, et al. The Gray Institute 'open' high-
612 content, fluorescence lifetime microscopes. *J Microsc*. 2013;251(2):154-67.

613 17. Barber PR, Ameer-Beg SM, Gilbey J, Carlin LM, Keppler M, Ng TC, et al. Multiphoton time-domain
614 FLIM: Practical application to protein-protein interactions using global analysis. *J R Soc Interface*. 2009;6:S93-
615 S105.

616 18. Rowley MI, Coolen ACC, Vojnovic B, Barber PR. Robust Bayesian Fluorescence Lifetime Estimation,
617 Decay Model Selection and Instrument Response Determination for Low-Intensity FLIM Imaging. *PLOS ONE*.
618 2016;11(6):e0158404.

619 19. Cravens PD, Hayashida K, Davis LS, Nanki T, Lipsky PE. Human peripheral blood dendritic cells and
620 monocyte subsets display similar chemokine receptor expression profiles with differential migratory responses.
621 *Scand J Immunol*. 2007;65(6):514-24.

622 20. Bronte V, Brandau S, Chen SH, Colombo MP, Frey AB, Greten TF, et al. Recommendations for myeloid-
623 derived suppressor cell nomenclature and characterization standards. *Nat Commun*. 2016;7:12150.

624 21. Nixon AB, Schalper KA, Jacobs I, Potluri S, Wang IM, Fleener C. Peripheral immune-based biomarkers in
625 cancer immunotherapy: can we realize their predictive potential? *J Immunother Cancer*. 2019;7(1):325.

626 22. Gatta G, Capocaccia R, Botta L, Mallone S, De Angelis R, Ardanaz E, et al. Burden and centralised
627 treatment in Europe of rare tumours: results of RARECAREnet-a population-based study. *Lancet Oncol*.
628 2017;18(8):1022-39.

629 23. Machiels JP, Rene Leemans C, Golusinski W, Grau C, Licitra L, Gregoire V, et al. Squamous cell
630 carcinoma of the oral cavity, larynx, oropharynx and hypopharynx: EHNS-ESMO-ESTRO Clinical Practice
631 Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2020;31(11):1462-75.

632 24. Cadoni G, Giraldi L, Petrelli L, Pandolfini M, Giuliani M, Paludetti G, et al. Prognostic factors in head and
633 neck cancer: a 10-year retrospective analysis in a single-institution in Italy. *Acta Otorhinolaryngol Ital*.
634 2017;37(6):458-66.

635 25. Ginhoux F, Mildner A, Gautier EL, Schlitzer A, Jakubzick C, Varol C, et al. Editorial: Monocyte
636 Heterogeneity and Function. *Front Immunol*. 2020;11:626725.

637 26. Zawada AM, Rogacev KS, Rotter B, Winter P, Marell RR, Fliser D, et al. SuperSAGE evidence for
638 CD14⁺⁺CD16⁺ monocytes as a third monocyte subset. *Blood*. 2011;118(12):e50-61.

639 27. Nenclares P, Gunn L, Soliman H, Bover M, Trinh A, Leslie I, et al. On-treatment immune prognostic score
640 for patients with relapsed and/or metastatic head and neck squamous cell carcinoma treated with immunotherapy. *J*
641 *Immunother Cancer*. 2021;9(6).

642

643