

# 1 External validation of prognostic and 2 predictive gene signatures in head and 3 neck cancer patients

4  
5 Erlend I. F. Fossen<sup>1,2,\*</sup>, Mauricio Moreira-Soares<sup>1</sup>, Marissa LeBlanc<sup>1,3</sup>, Arnaldo Frigessi<sup>1</sup>,  
6 Eivind Hovig<sup>4,5</sup>, Laura Lopez-Perez<sup>6</sup>, Estefanía Estévez-Priego<sup>6</sup>, Liss Hernandez<sup>6</sup>, Maria  
7 Fernanda Cabrera-Umpierrez<sup>6</sup>, Giuseppe Fico<sup>6</sup>, Ingeborg Tinhofer<sup>7</sup>, Vanessa Sachse<sup>7</sup>,  
8 Kathrin Scheckenbach<sup>8</sup>, Christophe Le Tourneau<sup>9</sup>, Maud Kamal<sup>9</sup>, Steve Thomas<sup>10</sup>, Miranda  
9 Pring<sup>10</sup>, Lisa Licitra<sup>11,12</sup>, Loris De Cecco<sup>13§</sup>, Stefano Cavalieri<sup>11,12§</sup>

10  
11 <sup>1</sup> Oslo Centre for Biostatistics and Epidemiology, Institute of Basic Medical Sciences,  
12 University of Oslo, Oslo, Norway

13 <sup>2</sup> CICERO Center for International Climate Research, Oslo, Norway

14 <sup>3</sup> Department of Method Development and Analytics, Norwegian Institute of Public Health,  
15 Oslo, Norway

16 <sup>4</sup> Centre for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway

17 <sup>5</sup> Department of Tumor biology, Institute for Cancer Research, Oslo University Hospital,  
18 0310 Oslo, Norway

19 <sup>6</sup> Universidad Politécnica de Madrid-Life Supporting Technologies Research Group, ETSIT,  
20 28040 Madrid, Spain

21 <sup>7</sup> Department of Radiooncology and Radiotherapy, Translational Radiation Oncology  
22 Research Laboratory, Charité – Universitätsmedizin Berlin, corporate member of Freie  
23 Universität Berlin and Humboldt Universität zu Berlin, Charitéplatz 1, 10117 Berlin,  
24 Germany

25 <sup>8</sup> Heinrich-Heine-Universität Düsseldorf, Dusseldorf, Germany

26 <sup>9</sup> Department of Drug Development and Innovation (D3i), Institut Curie, Paris-Saclay  
27 University, Paris, France

28 <sup>10</sup> Bristol Dental School, University of Bristol, United Kingdom

29 <sup>11</sup> Head and Neck Medical Oncology Department, Fondazione IRCCS Istituto Nazionale dei  
30 Tumori di Milano, via Giacomo Venezian 1, 20133 Milan, Italy

31 <sup>12</sup> Department of Oncology and Hemato-oncology, University of Milan, via Santa Sofia 9/1,  
32 20122 Milan, Italy

33 <sup>13</sup> Experimental Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, via Giacomo  
34 Venezian 1, 20133 Milan, Italy

35  
36 \*Corresponding author, [erlend.fossen@cicero.oslo.no](mailto:erlend.fossen@cicero.oslo.no)

37 § Stefano Cavalieri and Loris De Cecco equally contributed as co-last authors.

38

39

## 40 Abstract

41 Head and neck squamous cell carcinomas (HNSCC) are aggressive and heterogenous tumors  
42 with a high fatality rate. Many gene signatures (GS) have been developed with both  
43 prognostic and predictive significance. We aimed to externally validate five published GS in  
44 a large European collection of HNSCC patients. Gene expression from 1097 treatment-naïve  
45 HNSCC patients' primary tumors was used to calculate scores for the five GS. Cox  
46 proportional hazard models were used to test the association between both 2-year overall  
47 survival and 2-year disease-free survival and the signature scores. The predictive role of GS  
48 was validated by comparing survival associations in patients receiving specific treatment (i.e.,  
49 radiotherapy, systemic treatment) versus those who did not. We successfully externally  
50 validated all 5 GS, including two prognostic signatures, one signature as prognostic and  
51 predictive of sensitivity to systemic treatment, while signatures for cisplatin-sensitivity and  
52 radiosensitivity were validated as prognostic only.

53

## 54 1. Introduction

55 Head and neck cancer (HNC) is the 7<sup>th</sup> most common cancer worldwide, with more than  
56 890,000 new cases and 450,000 deaths annually<sup>1,2</sup>. Head and neck squamous cell carcinoma  
57 (HNSCC) is the most common histological type of HNC, making up 90% of all HNC  
58 cases<sup>2,3</sup>, roughly 4.5% of all global cancer diagnoses and approximately 4.6% of all cancer  
59 deaths<sup>1,2</sup>. Major risk factors for HNSCC are tobacco smoking and alcohol consumption<sup>1,3</sup>, but  
60 human papillomavirus (HPV) infection is also a known risk factor for oropharyngeal  
61 cancers<sup>4</sup>.

62

63 HNSCCs are a heterogenous group of cancers with malignancies developing at various  
64 anatomical sub-sites and with high molecular tumor heterogeneity<sup>5,6</sup>. This heterogeneity  
65 contributes to differences between patients in terms of treatment response with a need for  
66 personalized approaches<sup>6,7</sup>. In general, HNSCC patients with early disease (stage I-II) have  
67 high cure rates with single modality therapies (either surgery or radiotherapy). Subjects with  
68 loco-regionally advanced disease (stage III-IVa/b) are treated with multimodal approaches  
69 (i.e., various combinations of surgery, radiotherapy and systemic treatment)<sup>8</sup>. Despite optimal  
70 treatments, approximately half of patients develop recurrence within 2 years and half die with  
71 disease within 5 years from the diagnosis<sup>9</sup>. Reliable biomarkers and gene signatures (GS) are  
72 crucial for developing personalized treatment<sup>10,11</sup>. GS are sets of genes involved in a  
73 biological process that can provide information about the expected disease outcome  
74 (prognostic) and/or the response to a specific treatment (predictive)<sup>10,11</sup>. Prognostic signatures  
75 provide information about disease prognosis irrespective of therapy. Meanwhile, predictive  
76 signatures inform about how likely a patient is to respond to a specific therapy and can  
77 therefore enable more personalized treatment plans.

78

79 Many prognostic GS, calculated from gene expression in treatment-naïve tumors, have been  
80 developed for HNSCC<sup>12-16</sup>. External validation of already published GS is an important step  
81 towards incorporating GS in clinical practice<sup>17</sup>. In this study, we decided to consider only GS  
82 which were already published, and which were either disease- or treatment-specific. Among  
83 them is a 172-gene signature<sup>18</sup> (*172-GS*) that was developed to be prognostic of patient's risk

84 of relapse in HNSCC patients independently of HPV-status, but was later shown to not be  
85 prognostic of overall survival (OS) in HPV-positive patients<sup>19</sup>. However, in the same patient  
86 population, a specific three-cluster HPV signature<sup>20</sup> (*3 clusters HPV*) was developed and  
87 externally validated as prognostic of OS<sup>19</sup>. As for predictive signatures, the 10-gene  
88 radiosensitivity index<sup>21</sup> (*RSI*) is among the most validated pan-cancer GSs for prediction of  
89 sensitivity to radiotherapy<sup>22</sup>. Another pan-cancer predictive signature, in this case predictive  
90 of cisplatin-sensitivity<sup>23</sup> (*pancancer-cisplatin*), was recently developed, but has not been  
91 externally validated in HNSCC patients. Lastly, a signature that classifies HNC patients into  
92 six different subtypes/clusters was developed by De Cecco et al<sup>24</sup>. One of them (*Cl3-hypoxia*)  
93 showed hypoxic features and was found to be prognostic<sup>19,25</sup> and related to response to  
94 treatment with anti-EGFR agents (cetuximab and afatinib)<sup>26,27</sup>.

95  
96 Using a large collection of HNSCC patients with available gene expression data, we aimed to  
97 externally validate two prognostic gene signatures (*172-GS* and *3 clusters HPV*), two  
98 predictive gene signatures (*RSI* and *pancancer-cisplatin*) and one prognostic and predictive  
99 GS (*Cl3-hypoxia*). In contrast to earlier studies, we wanted to explicitly test whether the  
100 prognostic signatures (*172-GS* and *3 clusters HPV*) are valid independent of HPV-status, and  
101 explicitly test if the predictive signatures are predictive of treatment response by comparing if  
102 the signature effect is only found in patients with a specific treatment and not in patients  
103 without that treatment. We tested these signatures using two survival endpoints, OS and  
104 disease-free survival (DFS).  
105

## 106 **2. Methods**

### 107 **2.1. Ethics statement**

108 This study was conducted in full accordance with the World Medical Association's  
109 Declaration of Helsinki (2013 version). The protocol was ethically approved by the  
110 Norwegian Regional Committee for Medical and Health Research Ethics (REK) South-East  
111 A under application number 270467. The data is securely stored in the University of Oslo's  
112 server for sensitive data (TSD/USIT), adhering to the requirements of GDPR legislation.  
113 Access to the data is granted only to authorized collaborators who have been included in the  
114 ethical approval. All proprietary studies that contributed data obtained ethical approval from  
115 their respective local authorities in Italy, Germany, or France. Copies of these ethical  
116 approvals were provided to the principal investigator by the data providers. BD2Decide was  
117 approved by the Ethical Committee of the Fondazione IRCCS Istituto Nazionale dei Tumori  
118 (Milan, Italy) in 2016 and it has two identifiers: INT65-16 and INT66-16. Biomarker analysis  
119 of the ARO 04-01 Def-RCT cohort was approved by the Ethical Committee of the Charité -  
120 Universitätsmedizin Berlin (Berlin, Germany) in 2010 (EA2/086/10). Biomarker analysis of  
121 the DKTKRO Def-RCT was approved by the Ethical Committee of the Technical University  
122 of Dresden (Dresden, Germany) in 2014 (EK200112014).

123

### 124 **2.2. Patient selection**

125 Within a European cooperative research project named SuPerTreat (research grant nr.  
126 ERAPERMED2019-281, further details in Funding), we first identified and constructed a

127 multicenter dataset, which consisted primarily of data collected in previous research  
128 projects<sup>28-36</sup> (NCT02832102, NCT03017573, NCT02059668) at the Fondazione IRCCS  
129 Istituto Nazionale dei Tumori (Milan, Italy), Institut Curie (Paris, France) and Charité -  
130 Universitätsmedizin Berlin (Berlin, Germany) (Table S1 in supplementary information). We  
131 additionally included two public datasets (TCGA<sup>37,38</sup> and GSE41613<sup>39</sup>) that adhered to  
132 MIAME<sup>40</sup> and MINSEQE<sup>41</sup> standards.

133

134 The selected studies complied with the following **inclusion criteria**:

135

- 136 1. Diagnosis of head and neck squamous cell carcinoma (HNSCC)
- 137 2. Planned treatment with curative intent
- 138 3. Tumor specimens (formalin-fixed paraffin-embedded, FFPE) of HNSCC sampled  
139 prior to treatment, already assessed for gene expression analysis
- 140 4. Stage I, II, III, IVA or IVB according to the tumor-node-metastasis (TNM) cancer  
141 staging system 7th edition<sup>42</sup>
- 142 5. One of the following tumor sites: oral cavity, oropharynx, hypopharynx, larynx
- 143 6. Age  $\geq 18$  years

144 A total of N = 1589 patients were screened from these studies. From these, patients were  
145 excluded based on the following **general exclusion criteria**:

146

- 147 1. Gene expression data not available
- 148 2. Distant metastasis at diagnosis or missing information about stage
- 149 3. Recurrent/metastatic disease at the time of study entry
- 150 4. Missing information about survival timing or survival status

151

152 **Additional exclusion criteria** were applied when testing gene signatures related to  
153 sensitivity to systemic treatment:

- 154 5. Early disease stage excluded when testing signatures *pancancer-cisplatin* and *Cl3-*  
155 *hypoxia*. This specification is due to the fact that most early disease stage HNSCC  
156 patients are not treated with systemic treatment
- 157 6. Missing treatment information about systemic treatment agent when testing signatures  
158 *pancancer-cisplatin* and *Cl3-hypoxia*

159

160 A total of 1097 patients were eligible for inclusion in this study, with the number reduced for  
161 gene signatures related to sensitivity to systemic treatment (Figure 1, Table S1). In models  
162 testing the signatures *172-GS*, *3 clusters HPV* and *RSI*, all 1097 patients were eligible for  
163 analyzing OS (Figure 1) and 907 patients were eligible for DFS (Figure S1). In models  
164 testing sensitivity to systemic treatment signatures (*pancancer-cisplatin* and *Cl3-hypoxia*),  
165 750 patients were eligible for analyses of OS (Figure S2) and 700 patients were eligible for  
166 analyses of DFS (Figure S3).

167

168 None of the patients that were eligible for testing the signatures *172-GS*, *RSI*, *pancancer-*  
169 *cisplatin* or *Cl3-hypoxia* were previously included in the development of these signatures. For  
170 the *3 clusters HPV* GS, most of the eligible HPV-positive patients (N = 152, 14% of all  
171 eligible patients) were already part of a previous external validation of the signature<sup>19</sup>, but in  
172 this study we additionally tested the GS in HPV-negative patients and adjusted for more  
173 covariates.

174

## 175 **2.3. Endpoints, clinical variables and data harmonization**

176 OS, defined as survival time from diagnosis until death of any cause or censoring, was  
177 recorded in all studies. DFS, defined as time from diagnosis until cancer recurrence or death  
178 of any cause, whichever occurred first, or censoring, was recorded in all but two studies  
179 (Table S1). Patients were followed for up to 20 years in the period 1989-2017. Co-primary  
180 endpoints were 2-year OS and 2-year DFS.

181  
182 GSs were considered prognostic if they were associated with general disease prognosis  
183 independently of treatments. We defined a GS as predictive if the prognostic capability was  
184 limited to a population receiving a specific treatment.

185  
186 Data harmonization of clinical variables was performed following head and neck ontology  
187 (HeNeCON<sup>43</sup>) and quality rules defined in BD2Decide study<sup>28</sup> (extended in the context of the  
188 SuPerTreat project), to achieve comparable measurements across studies (see Table 1 and  
189 Tables S2 and S3 for full list of clinical covariates). Tumor staging at diagnosis was  
190 performed according to the 7th edition of the AJCC/UICC staging system<sup>42</sup>. The staging  
191 variable was then classified in two main groups based on disease extension: early disease  
192 (stage I-II) vs. loco-regionally advanced disease (stage III-IVa/b). Human papillomavirus  
193 (HPV) infection is a known risk factor for oropharyngeal cancer<sup>4</sup> and HPV-status was  
194 recorded for tumor tissues of all patients with oropharyngeal cancer, using either HPV  
195 DNA/RNA tests (with DNA PCR or ISH with or without E6/E7 mRNA confirmation, as per  
196 local guidelines) or through p16ink4a immunostaining. Non-oropharyngeal HNSCCs  
197 (including oral cavity cancers<sup>44</sup>) were considered HPV-negative in analyses. Systemic  
198 treatments were coded in two ways: 1) platinum-based (if carboplatin and/or cisplatin were  
199 part of the agents a patient received) or non-platinum based (primarily cetuximab or  
200 paclitaxel); and 2) cetuximab-based (if cetuximab was part of the agents a patient received) or  
201 non-cetuximab based. Patients not receiving systemic treatments were in both cases coded as  
202 such.

203  
204 Treatments (surgery, radiotherapy and systemic therapy) performed with curative intent were  
205 recorded as received or not received. This is potentially problematic since Cox regression  
206 survival models assume that covariates are measured at baseline (diagnosis) or require  
207 information about the timing of treatment. The use of the received treatment as a proxy for  
208 intended treatment (a baseline covariate) may result in immortal time bias<sup>45</sup>. We performed  
209 extensive sensitivity analyses to evaluate if our results were sensitive to using received  
210 treatment as a covariate (Appendix 1). First, we considered the degree of mortality observed  
211 during the period prior to when treatments were initiated. Second, we used an external dataset  
212 (Head and Neck 5000<sup>46,47</sup>) with similar patient characteristics as SuPerTreat and compared  
213 treatment coefficients from Cox models with either intended treatment or received treatment  
214 as covariates. Third, we compared treatment coefficients from Cox models with received  
215 treatment with models where treatment was considered time-dependent (in a subset of  
216 patients where the timing of surgery was known). Fourth, we used a landmark analysis<sup>48,49</sup>  
217 and compared treatment coefficients from a Cox model with received treatment with a  
218 landmark model where patients without an event or censoring prior to a landmark time point  
219 (corresponding with a time at which all patients started treatment) were followed from the  
220 landmark time. Lastly, we used quantitative bias analyses<sup>50</sup> to compare estimates of received  
221 treatments with bias-adjusted estimates of treatments. Based on these sensitivity analyses  
222 (Results section 3.3.1), we inferred that our models are robust to violations of the  
223 assumptions of the Cox regression models, and therefore used received treatment as a

224 covariate in our survival models. Although gene signatures are tested in interaction with  
225 treatments, the tumor tissue used to obtain gene signature scores were sampled at diagnosis  
226 (prior to treatment).  
227

## 228 2.4. Bioinformatics analysis

229 Gene expression data was profiled on different platforms depending on the study. Three  
230 studies (study IDs: BD2\_INT\_MI, BD2\_UDUS, GSE41613) were profiled on Affymetrix  
231 microarrays (platforms GPL23126 and GPL570), one study (GECO110) was profiled on  
232 Illumina microarray (platform GPL14951), and two studies (SCANDARE, TCGA) were  
233 profiled using Illumina RNA-seq. Within studies, Affymetrix data were quantile-normalized  
234 using robust multi-array average (RMA). Illumina microarray data were quantile-normalized.  
235 For RNA-seq data, within studies, we first removed very lowly expressed genes, where genes  
236 were removed if all samples had < 0.25 counts per million (cpm), corresponding to max 5  
237 reads in the smallest sample and max 22 reads in the largest sample. RNA-seq data were then  
238 quantile-normalized using the limma R package<sup>51</sup> (function: voom). We used WGCNA  
239 package<sup>52</sup> (function collapseRows and “maxRowVariance” method) to select the most  
240 variable probe in cases when multiple probes mapped to the same EntrezID. Probes and  
241 ensemble IDs that mapped to multiple EntrezID were removed. All gene expression data  
242 were log<sub>2</sub>-transformed after the abovementioned normalization. Genes that were missing for  
243 all patients in a study were removed when combining datasets. ComBat<sup>53,54</sup> was used to  
244 remove batch effects introduced by systemic non-biological technical errors from combining  
245 data from multiple studies with different platforms. The largest proprietary study (ID:  
246 BD2\_INT\_MI) was used as reference in ComBat, parametric adjustment was performed, and  
247 since there is variability in patient characteristics between studies, we only adjusted for the  
248 mean to not remove biological variance between the studies. To avoid bias in downstream  
249 analyses<sup>55</sup>, we only included the batches (studies) as covariates in ComBat.

250  
251 We calculated 5 gene signature scores for each patient following descriptions given in the  
252 original publications (see Table S4 for detailed list of genes and weights per signature). The  
253 signatures *172-GS*<sup>18</sup>, *3 clusters HPV*<sup>19,20</sup>, *pancancer-cisplatin*<sup>23</sup>, and *Cl3-hypoxia*<sup>24</sup> were  
254 calculated as weighed sums using the equation:  $Gene\ signature\ score = \sum_{i=1}^n \beta_i \times Exp_i$ ,  
255 where  $\beta_i$  is the weight of the *i*-th gene (of *n* genes) and  $Exp_i$  is the gene expression of the *i*-th  
256 gene. The *172-GS* signature was calculated using the hacksig package<sup>56</sup>. The radiosensitivity  
257 index (*RSI*) signature is based on the expression of 10 specific genes. *RSI* was calculated as a  
258 weighed ranked sum using the equation:  $RSI = \sum_{i=1}^{10} \beta_i \times Ranked\_Exp_i$ , where  $\beta_i$  is the  
259 weight of the *i*-th gene and  $Ranked\_Exp_i$  is the ranked gene expression of the *i*-th gene. For  
260 each patient the 10 genes were ranked in ascending order using patient-specific expression  
261 data, where the gene with the lowest expression is assigned a value of 1, and the gene with  
262 the highest expression is assigned a value of 10.

263  
264 Each gene signature was mean-centered and scaled by the standard deviation (SD) within its  
265 respective dataset to make the scores easier to interpret (mean = 0, SD = 1, see Appendix 2  
266 for means and SD on original scale). In our OS dataset, 76% of the *172-GS* signature genes  
267 were present, 46% of the *3 clusters HPV* signature genes were present, all *RSI* genes were  
268 present, 92% of the *pancancer-cisplatin* genes were present, and 94% of the *Cl3-hypoxia*  
269 genes were present. For our DFS dataset, the presence of genes was 80% for the *172-GS*  
270 signature, 57% for the *3 clusters HPV* signature, all *RSI* genes present, 92% for *pancancer-*

271 *cisplatin* and 94% for *Cl3-hypoxia* genes (see Appendix 2 for a list of missing genes per gene  
272 signature). Any missing genes were missing for all individuals, making regression-based  
273 imputation not a viable option. Instead, all missing genes were imputed as having zero  
274 expression. Note that the scaled signature values do not change if we instead were to impute  
275 with another constant (e.g., imputing the overall mean expression).  
276

## 277 **2.5. Statistical analyses**

278 All statistical analyses were performed in R Statistical Software<sup>57</sup> (v4.1.3).  
279

### 280 **2.5.1. Handling of missing clinical covariates**

281 There were missing data in several categorical variables: smoking status, tumor stage and  
282 received treatments (i.e., surgery, systemic therapy and radiotherapy). The missingness was  
283 largely structural, resulting from some variables not being measured in some of the studies,  
284 and therefore regression-based imputation of missing data was not a viable option<sup>58</sup>. Instead,  
285 we used a missing indicator where missing values were coded as belonging to a “missing”  
286 category (see Table 1 for amounts of missingness).  
287

### 288 **2.5.2. Clinical base models**

289 Clinical base models, i.e., those not including the gene expression data, were built as  
290 benchmark models for models which included gene signatures as covariates. Four Cox  
291 proportional hazard models were fit with either 2-year OS or 2-year DFS as endpoints. Two  
292 models were fit with the same data used to test signatures not related to sensitivity to  
293 systemic treatment (*172-GS*, *3 clusters HPV* and *RSI*) and two models were fit with the same  
294 data used to test signatures related to sensitivity to platinum-based or cetuximab-based  
295 treatments (*pancancer-cisplatin* and *Cl3-hypoxia*, respectively). In all models, age, sex,  
296 tumor region and HPV-status, smoking status, undergone surgery (yes/no), received  
297 radiotherapy (yes/no), and study ID were used as covariates. In models testing signatures not  
298 related to systemic treatment, disease extension (early stage, locoregionally advanced) and  
299 received systemic therapy (yes/no) were also used as covariates. In models testing signatures  
300 related to systemic treatment, TNM stage and a variable for systemic treatment (platinum-  
301 based version when testing *pancancer-cisplatin*; cetuximab-based version when testing *Cl3-*  
302 *hypoxia*) were included as covariates. The concordance index (Harrell’s C-index<sup>59</sup>) was  
303 calculated for each model to compare their discriminatory ability with the ability of models  
304 that included gene signatures (but see Hartman et al.<sup>60</sup> for limitations on interpreting the C-  
305 index). Additionally, we calculated each models measure of explained variation ( $R^2$ ) as  
306 defined by Royston<sup>61</sup> and building on previous work<sup>62,63</sup>. This measure is similar to  $R^2$  of  
307 linear models but was made for models analyzing censored survival data<sup>61</sup>.

### 308 **2.5.3. Gene signature validation models**

309 To validate prognostic gene signatures (*172-GS* and *3 clusters HPV*), we fit Cox proportional  
310 hazard models where we tested if the gene signature was associated with survival endpoints  
311 while adjusting for other covariates. For each GS, we used the same endpoints and covariates

312 as in the clinical base models, but also tested for an interaction between HPV-status and the  
313 GS. This allowed us to test if the same association between survival and GSs was found in  
314 both HPV-positive and in HPV-negative patients.

315  
316 To validate predictive gene signatures (*RSI*, *pancancer-cisplatin* and *Cl3-hypoxia*), we fit  
317 Cox proportional hazard models where we tested if gene signatures were associated with  
318 survival endpoints and if there was evidence that gene signatures modified the effect of the  
319 treatments (i.e., testing for an interaction between gene signatures and treatments). Finding a  
320 significant interaction where a GS is associated with survival in a specific treatment but not  
321 in others would imply that the signature is predictive. For validating each GS, we used the  
322 same covariates as in the clinical base models, but also included relevant interactions between  
323 GS and treatments. An interaction between radiotherapy and *RSI* was included when  
324 validating *RSI*. For validation of the *pancancer-cisplatin* signature, we included an  
325 interaction between the *pancancer-cisplatin* signature and systemic treatment. Similarly for  
326 validation of *Cl3-hypoxia*, we included an interaction between the *Cl3-hypoxia* signature and  
327 systemic treatment (see Appendix 3 for formulas for each GS model).

328  
329 To test the proportional hazard assumption of our Cox models, we performed sensitivity  
330 analyses where we fit equivalent models as described above, using 5-year survival (both for  
331 OS and DFS) as the endpoint instead of 2-year survival. We additionally estimated the  
332 median follow-up time using the reverse Kaplan-Meier method<sup>64</sup>.  
333

## 334 **3. Results**

### 335 **3.1. Cohort characteristics**

336 The largest dataset used for overall survival analyses, representing the full SuPerTreat cohort,  
337 contained patients from 6 studies, where 67% of patients were from the two largest studies  
338 (Table 1). The majority of patients were men (72%) and the cohort had a median age at  
339 diagnosis of 61 years (SD = 12). Most patients had loco-regionally advanced disease (85%)  
340 and underwent surgery (78%) and/or radiotherapy treatment (69%), with 42% of patients  
341 receiving systemic treatment. There was heterogeneity in tumor regions, and the most  
342 frequent primary site was oral cavity (55%) (Table 1). Similar cohort characteristics were  
343 observed in the dataset used for disease-free survival (Table 1), and in subsets of patients that  
344 were used to test the effect of other GSs on survival endpoints (Tables S2, S3). Among  
345 patients who received systemic treatment with information available about the therapeutic  
346 agent, 87.6% received platinum-based and 8.5% received cetuximab-based therapies (Table  
347 S3).

### 348 **3.2. Gene signatures**

#### 349 **3.2.1. Prognostic signature: 172-GS**

350 The *172-GS* signature was significantly associated with 2-year OS in HPV-negative patients,  
351 where increasing the signature score by 1 (representing an increase of 1 standard deviation)  
352 resulted in 41% higher hazard (HR = 1.41 [CI: 1.25, 1.59], Table 2, Figure 2). No evidence of  
353 an interaction between the signature score and HPV-status was found ( $p = 0.621$ ), indicating  
354 that the association between survival and the GS was similar in HPV-positive patients (HR =  
355 1.87 [CI: 0.61, 5.78], Table 2). However, the score was not significantly associated with OS



356 in HPV-positive patients when testing the marginal effect of the signature conditioned on  
357 being HPV-positive (Table 2). When fixing the signature score at zero (representing the mean  
358 score), HPV-negative patients had a 4.27-fold higher hazard than HPV-positive patients  
359 (Table 2). The models C-index was 0.71 (SE = 0.02) and  $R^2 = 0.29$ . The C-index of the  
360 corresponding clinical base model (i.e. not including the GS as a covariate) was 0.69 (SE =  
361 0.02) and  $R^2 = 0.23$ .

362

363 Similar effect sizes and significance levels were found for 2-year DFS, with a significant  
364 association between the signature score and survival in HPV-negative patients (HR = 1.36  
365 [CI: 1.17, 1.59], Table 2). There was no evidence of an interaction between the signature  
366 score and HPV-status ( $p = 0.651$ ), and the score was not significantly associated with DFS in  
367 HPV-positive patients when conditioned on being HPV-positive (Table 2). The models C-  
368 index was 0.67 (SE = 0.02) and  $R^2 = 0.22$ . The corresponding clinical base model had a C-  
369 index of 0.66 (SE = 0.02) and  $R^2 = 0.18$ . See Table S5 for details about other covariates  
370 included in the *172-GS* models.

371

### 372 **3.2.2. Prognostic signature: 3 clusters HPV**

373 There was a significant association between the *3 clusters HPV* signature and 2-year OS in  
374 HPV-positive patients, where increasing the signature score resulted in a lower hazard (HR =  
375 0.41 [CI: 0.24, 0.71], Table 2, Figure 3). There was evidence of an interaction between the  
376 signature score and HPV-status ( $p = 0.007$ ), indicating that the effect of the GS was different  
377 in HPV-negative patients (HR = 0.89 [CI: 0.77, 1.02], Table 2). Moreover, the score was not  
378 significantly associated with OS in HPV-negative patients (Table 2). These patients had a  
379 4.17-fold higher hazard than HPV-positive patients when fixing the signature score at zero  
380 (Table 2). The model had a  $R^2 = 0.25$  and C-index = 0.69 (SE = 0.02). The C-index of the  
381 corresponding clinical base model was 0.69 (SE = 0.02) and  $R^2 = 0.23$ .

382

383 We found similar effect sizes and significance levels for 2-year DFS, with a significant  
384 association between the GS and survival in HPV-positive patients (HR = 0.51 [CI: 0.33,  
385 0.77], Table 2). Evidence of an interaction between the signature score and HPV-status was  
386 also found ( $p = 0.001$ ), and the score was not significantly associated with DFS in HPV-  
387 negative patients (Table 2). The models C-index was 0.67 (SE = 0.02) and  $R^2 = 0.20$ . The C-  
388 index of the corresponding clinical base model was 0.66 (SE = 0.02) and  $R^2 = 0.18$ . See  
389 Table S6 for details about other covariates included in the *3 clusters HPV* models.

390

### 391 **3.2.3. Predictive signature: radiosensitivity index (RSI)**

392 The *RSI* was significantly associated with 2-year OS in patients receiving radiotherapy,  
393 where a higher hazard was found when increasing the signature score (HR = 1.25 [CI: 1.08,  
394 1.44], Table 2, Figure 4). We found no evidence of an interaction between *RSI* and  
395 radiotherapy ( $p = 0.672$ ), suggesting that the association between survival and *RSI* was the  
396 same, although not significantly associated, in patients that did not receive radiotherapy (HR  
397 = 1.17 [0.90, 1.52], Table 2, Figure 4). When *RSI* was set to zero, there was no significant  
398 association between survival and having received radiotherapy (Table 2). The models C-  
399 index was 0.70 (SE = 0.02) and  $R^2 = 0.25$ . The corresponding clinical base model had a C-  
400 index of 0.69 (SE = 0.02) and  $R^2 = 0.23$ .

401

402 In contrast to OS, there was no significant association between 2-year DFS and *RSI* (Table 2).  
403 There was also no evidence of an interaction between *RSI* and having received radiotherapy  
404 ( $p = 0.601$ ). The models C-index was 0.66 (SE = 0.02) and  $R^2 = 0.18$ . The corresponding  
405 clinical base models C-index was 0.66 (SE = 0.02) and  $R^2 = 0.18$ . See Table S7 for details  
406 about other covariates included in the *RSI* models.  
407

#### 408 **3.2.4. Predictive signature: *Pancancer-cisplatin***

409 There was a significant association between the *pancancer-cisplatin* signature and 2-year OS  
410 in patients receiving platinum-based chemotherapy, where a higher hazard was observed  
411 when increasing the signature (HR = 1.29 [CI: 1.04, 1.59], Table 2, Figure 5). When  
412 platinum-based chemotherapy was the reference, there was no evidence of interactions  
413 between the signature and systemic treatment ( $p = 0.200$  when compared to non-platinum  
414 based,  $p = 0.908$  when compared to no systemic therapy), suggesting that the association  
415 between the signature and overall survival may be similar in these groups. However, there  
416 was no significant association between survival and the signature score when conditioning on  
417 receiving non-platinum based systemic therapy or when not receiving systemic treatment  
418 (Table 2, Figure 5). There was also no significant association between survival and systemic  
419 treatment when the signature score was set to zero (Table 2). The models  $R^2 = 0.32$  and C-  
420 index was 0.72 (SE = 0.02). The corresponding clinical base model had a C-index of 0.71  
421 (SE = 0.02) and  $R^2 = 0.30$ .

422  
423 In contrast to OS, we found no significant association between 2-year DFS and the  
424 *pancancer-cisplatin* signature (Table 2). There was also no evidence of interactions between  
425 the score and systemic treatment when platinum-based chemotherapy was the reference ( $p =$   
426  $0.258$  when compared to non-platinum based,  $p = 0.571$  when compared to no systemic  
427 treatment). The models C-index was 0.70 (SE = 0.02) and  $R^2 = 0.28$ . The C-index of the  
428 corresponding clinical base model was 0.70 (SE = 0.02) and  $R^2 = 0.26$ . See Table S8 for  
429 details about other covariates included in the *pancancer-cisplatin* models.  
430

#### 431 **3.2.5. Predictive and prognostic signature: *Cl3-hypoxia***

432 The *Cl3-hypoxia* signature was significantly associated with 2-year OS in patients receiving  
433 non-cetuximab-based systemic treatments, where increasing the signature score resulted in a  
434 lower hazard (HR = 0.69 [CI: 0.54, 0.87], Table 2, Figure 6). A similar, but non-significant,  
435 association between OS and the GS was found for cetuximab-treated patients (HR = 0.67 [CI:  
436 0.39, 1.14], Table 2, Figure 6). In contrast, an opposite association was found between OS  
437 and the signature in patients who did not receive systemic treatments, where higher signature  
438 scores were associated with higher hazards (HR = 1.70 [CI: 1.12, 2.58], Table 2, Figure 6).  
439 There was evidence of interactions between the signature and systemic treatment when using  
440 no systemic treatment as the baseline (comparison with cetuximab group:  $p = 0.006$ ;  
441 comparison with non-cetuximab:  $p < 0.001$ ), but no evidence of an interaction when  
442 comparing cetuximab-treated patients with patients receiving other types of systemic  
443 treatment ( $p = 0.919$ ). When the GS score was set to zero, there was no significant  
444 association between survival and systemic treatment (Table 2). The models C-index was 0.73  
445 (SE = 0.02) and  $R^2 = 0.34$ . The C-index of the corresponding clinical base model (i.e. not  
446 including the GS as a covariate) was 0.71 (SE = 0.02) and  $R^2 = 0.29$ .

447

448 Similar effect sizes and significance levels were found when analyzing 2-year DFS (Table 2).  
449 Increasing the signature score was associated with a reduced hazard in patients receiving  
450 systemic treatment, while being significantly associated with an increased hazard in patients  
451 not receiving systemic treatments (Table 2). There was evidence of interactions between the  
452 signature and systemic treatment when no systemic treatment was the baseline (comparison  
453 with cetuximab group:  $p = 0.006$ ; comparison with non-cetuximab:  $p < 0.001$ ), but no  
454 evidence of an interaction when comparing patients receiving cetuximab with patients  
455 receiving non-cetuximab therapies ( $p = 0.943$ ). The models  $R^2 = 0.32$  and C-index = 0.71 (SE  
456 = 0.02). The C-index of the corresponding clinical base model was 0.70 (SE = 0.02) and  $R^2 =$   
457 0.27. See Table S9 for details about other covariates included in the *Cl3-hypoxia GS*.  
458

### 459 **3.3. Sensitivity analyses**

#### 460 **3.3.1. Coding of treatment**

461 We performed sensitivity analyses to evaluate the effect of using the received treatment as a  
462 proxy of intended treatment (Appendix 1). Results of these analyses are summarized below:

- 463 i) All treated patients started treatment within 6 months after diagnosis, with most  
464 starting treatment within 3-4 months. In the first 6 months after diagnosis, the  
465 mortality was approximately 3%.
- 466 ii) We obtained overlapping and similar treatment coefficients when comparing  
467 intended and received treatment in an external dataset with similar patients'  
468 characteristics.
- 469 iii) We obtained overlapping and similar treatment coefficients when comparing  
470 received treatment with time-dependent treatment in a subset of patients where the  
471 timing of surgery was known.
- 472 iv) A landmark analysis showed overlapping and similar treatment coefficients as  
473 when using received treatment.
- 474 v) Quantitative bias analyses showed bias-adjusted estimates of treatment effects that  
475 overlap with estimates obtained when using received treatment.  
476

#### 477 **3.3.2. 5-year endpoints**

478 Effect sizes and significance levels from models with 5-year OS and 5-year DFS were very  
479 similar to the 2-year survival models (see Figures S4-S13 for forest plots and Tables S5-S9  
480 for detailed results). The main difference was in the model of 5-year OS and the *Cl3-hypoxia*  
481 signature, where there was no longer evidence of an association between the score and  
482 survival in patients not receiving systemic treatments (HR = 1.43 [CI: 0.96, 2.13],  $p = 0.076$ ).  
483 Moreover, there was no longer evidence of an interaction between the score and systemic  
484 treatment when comparing patients untreated with systemic agents with patients receiving  
485 cetuximab ( $p = 0.068$ ). The median follow-up time was 1325 days (approx. 3.63 years).  
486

## 487 **4. Discussion**

488 In this study, we externally validated five gene signatures using one of the world's largest  
489 collections of HNSCC patients with available gene expression, harmonizing and combining

490 high quality clinical data from different studies. The development of this wide European  
491 dataset has shown that the integration of multisource clinical and biological databases is  
492 feasible. This successful integration is unprecedented and amplifies the knowledge about the  
493 complexity and the heterogeneity of HNSCCs, contributing to their qualification and  
494 quantification. The external validation ensures the robustness of the GSs and their potential  
495 generalizability. In addition, tumor heterogeneity and complexity coupled with the variety of  
496 treatment options imposes a biology-driven personalized approach in the curative setting of  
497 HNSCC patients. Unlike more frequent tumors (e.g., breast cancer, non-small cell lung  
498 cancer, or colorectal cancer), GSs have not been included yet either in clinical decision  
499 making, or in clinical trial eligibility and stratification of HNSCC patients.

500  
501 Except for the model testing associations between *RSI* and disease-free survival, all models  
502 including gene signatures outperformed their corresponding clinical omics-free base models,  
503 as indicated by C-index and/or  $R^2$ . Our results confirmed two potentially prognostic  
504 signatures (*172-GS* and *3 clusters HPV*) and validated one potentially prognostic/predictive  
505 signature (*Cl3-hypoxia*) for cetuximab- and chemosensitivity. However, the results for two  
506 potentially predictive signatures (*RSI* and *pancancer-cisplatin*) were less conclusive, despite  
507 showing associations with overall survival in patients receiving radiotherapy or platinum-  
508 based chemotherapy.

509  
510 The signature *172-GS*<sup>18</sup> was validated as a prognostic indicator for both overall and disease-  
511 free survival in patients with HPV-unrelated disease. A higher signature score is associated  
512 with a higher risk. The signature may also be prognostic in HPV-positive oropharyngeal  
513 cancer patients, as given by the lack of evidence of an interaction between the signature and  
514 HPV-status. This contrasts with previously published results<sup>19</sup>, possibly due to adjustments  
515 for covariates and uncertainties in estimates of the signature effect for HPV-positive patients.

516  
517 The *3 clusters HPV* signature<sup>20</sup> was validated as prognostic for both OS and DFS in HPV-  
518 positive oropharyngeal cancer patients, consistent with a previous validation study<sup>19</sup>.  
519 Additionally, there was no evidence of the signature being prognostic for HPV-negative  
520 patients, suggesting its relevance specifically to HPV-positive oropharyngeal cancer.

521  
522 In our study, we found a link between the radiosensitivity index (*RSI*)<sup>21</sup> and overall survival  
523 in patients who underwent radiotherapy, consistent with other studies<sup>18,19,22,65</sup>. However, we  
524 did not observe any interaction between *RSI* and radiotherapy. We also found a similar but  
525 not significant association between *RSI* and OS in patients who did not receive radiotherapy.  
526 This suggests that *RSI* may be more prognostic for survival than predictive of  
527 radiosensitivity. This surprising finding of a prognostic effect of *RSI* in non-irradiated  
528 patients may be because of the relatively low number of non-irradiated patients in our study.  
529 In contrast to other studies<sup>22,66</sup>, we did not find a connection between *RSI* and DFS.

530  
531 To our knowledge this is the first external validation of the *pancancer-cisplatin* signature<sup>23</sup> in  
532 head and neck cancer patients. We found a negative association between the signature score  
533 and overall survival (OS) in patients receiving platinum-based chemotherapy. Given that  
534 higher GS scores were linked to higher cisplatin-sensitivity in the original publication<sup>23</sup>, our  
535 finding is in the opposite direction of what is expected. Additionally, we did not find  
536 evidence of an interaction between systemic treatment and the signature. Due to an imbalance  
537 in the number of patients with each type of systemic treatment, we could not draw definitive  
538 conclusions about the potential prognostic or predictive effects of the signature. Lastly, we  
539 found no associations between disease-free survival (DFS) and the *pancancer-cisplatin*

540 signature, suggesting that its predictive power may vary with clinical outcomes and not be  
541 universal for all cancers.

542

543 In general, a GS is prognostic if it is able to forecast patient survival independently of the  
544 received treatment. At the same time, if a prognostic GS is tested on a subset of patients  
545 receiving curative therapies including a specific approach, then the prognostic performance  
546 necessarily implies a certain predictive capability. In this scenario, the *Cl3-hypoxia*<sup>24</sup>  
547 signature score was positively associated with both OS and DFS in patients receiving  
548 systemic treatment but negatively associated with survival in subjects not receiving systemic  
549 agents, suggesting the signature is predictive of sensitivity to systemic therapy. A similar  
550 positive association with survival endpoints was found in both cetuximab-treated patients and  
551 in patients with chemotherapy, which did not include anti-EGFR agents. The association  
552 between signature and survival is consistent with previous findings in HPV-positive  
553 oropharyngeal cancer<sup>19</sup> and in oral premalignant lesions<sup>25</sup>. Previous studies have suggested  
554 that the signature is related to cetuximab-sensitivity<sup>26,27</sup>, but our results indicate that it is  
555 predictive of sensitivity to systemic treatment in general.

556

557 This study offers various opportunities for further improvement and refinement. Firstly, we  
558 used the received treatment as a proxy for intended treatment. While this breaks an  
559 assumption of the Cox regression model, we performed extensive sensitivity analyses that  
560 suggest the models are robust to this violation. Secondly, 3-7% of patients (depending on the  
561 model) were excluded for missing overall survival status or timing, and 6-19% were excluded  
562 for missing DFS. This complete case exclusion reduced the sample sizes. Still, because the  
563 missingness was primarily structural<sup>58</sup>, where the survival endpoint was missing for all  
564 patients in a study, it is unlikely that it biased the results. Third, we used the missing indicator  
565 method for missing covariates (with missingness ranging from 3-30% depending on the  
566 covariate), which can lead to biased results in non-randomized studies<sup>67</sup>, but is unlikely to  
567 result in bias unless the covariate is a strong confounder or missing in extreme proportion  
568 (>50% missing)<sup>68</sup>. Moreover, like the missingness in endpoints, the missingness in covariates  
569 was primarily structural, and alternative regression-based imputation methods were therefore  
570 not a viable option. Lastly, the imbalance in the number of patients with different treatments  
571 (e.g., only 35 patients received cetuximab) made it difficult to precisely estimate interactions  
572 and effects of gene signatures in small groups and made conclusions less clear.

573

574 We externally validated five gene signatures using a large integrated dataset of HNSCC  
575 patients. Our results validated two prognostic signatures, *172-GS* and *3 clusters HPV*  
576 signature, in HPV-negative patients and HPV-positive patients, respectively. We also  
577 validated that the potentially predictive signatures *RSI* and *pancancer-cisplatin* were  
578 prognostic of survival in patients with radiotherapy and platinum-based chemotherapy,  
579 respectively, but could not conclude if these signatures are also predictive of radiosensitivity  
580 or platinum-chemosensitivity. Lastly, we validated the *Cl3-hypoxia* signature as predictive of  
581 sensitivity to systemic treatment, making this signature a good candidate for use in  
582 personalized treatment decisions.

583

584 Many clinical studies aimed at intensifying<sup>69-71</sup> or de-escalating<sup>72-74</sup> standard treatments have  
585 failed in head and neck oncology. A possible drawback could be the lack of an appropriate  
586 patient selection, which may have hampered the opportunity to detect a signal of clinical  
587 benefit with a given therapeutic approach. The significant biological differences in the  
588 current study demonstrate the high heterogeneity of HNSCCs. Therefore, future clinical trials  
589 should pursue a better patient selection that includes a broader use of GS and incorporates

590 them as inclusion criteria or stratification factors. In this scenario, the results of our external  
591 validation in one of the widest HNSCC datasets may be considered a historical benchmark  
592 for the design of future studies.  
593

## 594 **5. Data availability**

595 The SuPerTreat dataset is hosted by the Services for Sensitive Data (TSD) at the University  
596 of Oslo. Anonymized data containing only the survival endpoints and gene signature scores  
597 will be uploaded to Zenodo prior to publication. Access to the original data may be granted  
598 by the data owners upon application.

## 599 **6. Code availability**

600 The underlying code (R scripts) used to analyse and validate gene signatures is available on  
601 GitHub and can be accessed via this link  
602 [https://github.com/erlendfossen/SuPerTreat\\_GS\\_validation](https://github.com/erlendfossen/SuPerTreat_GS_validation).  
603

## 604 **7. Acknowledgements**

605 This work is part of the research project “Supporting Personalized Treatment Decisions in  
606 Head and Neck Cancer through Big Data (SuPerTreat)” funded within ERA PerMed  
607 JTC2019 Joint transnational call for proposals (2019) for “Personalised medicine:  
608 multidisciplinary research Toward implementation”, research grant nr. ERAPERMED2019-  
609 281 and supported by FRRB (Fondazione Regionale per la Ricerca Biomedica), BMBF  
610 (Bundesministerium für Bildung und Forschung), RCN (Norges forskningsråd), ANR  
611 (Agence nationale de la recherche) and GSRT (General Secretariat for Research and  
612 Technology). MM-S received funding from the European Union’s Horizon 2020  
613 Research and Innovation program under the Marie Skłodowska-Curie Actions Grant,  
614 agreement No. 80113 (Scientia fellowship). BD2Decide received funding from the European  
615 Union Horizon 2020 Framework Programme, Grant/Award Number: 689715. We thank the  
616 researchers, clinicians and other staff that designed and conducted the studies that this work  
617 builds on. We also thank the head and neck cancer patients that took part in the studies.  
618

## 619 **8. Author contributions**

620 Conceptualization: E.I.F.F., M.M.S., M.L., A.F., I.T., C.L.T., L.L, L.D.C, S.C.; Data curation:  
621 E.I.F.F., M.M.S., L.L.P, E.E.P, L.H, L.D.C, S.C.; Formal analysis: E.I.F.F., M.M.S., M.L.,  
622 A.F., L.D.C, S.C.; Funding acquisition: A.F., E.H., M.F.C.U, G.F., I.T., L.L.; Investigation:  
623 I.T., V.S., K.S., C.L.T., M.K., S.T., M.P., L.L, L.D.C, S.C.; Methodology: E.I.F.F., M.M.S.,  
624 M.L., A.F., L.L, L.D.C, S.C.; Resources: A.F., L.L.; Software: E.I.F.F., M.M.S.; Supervision:  
625 M.L., A.F., M.F.C.U., G.F.; Visualization: E.I.F.F.; Writing-original draft: E.I.F.F.; Writing-

626 review & editing: All authors. All authors approved the submission of this manuscript. L.D.C.  
627 and S.C. contributed equally to this work as co-last authors.

628

## 629 **9. Competing interests**

630 M.L. reports receiving a speaker fee from MSD unrelated to the content of this work. L.L.  
631 declares the following conflicts of interest, all unrelated to the content of this work: research  
632 funds donated directly to the institute for clinical trials from AstraZeneca, BMS, Boehringer  
633 Ingelheim, Celgene International, Eisai, Exelixis, Debiopharm International SA, Hoffmann-  
634 La Roche Ltd, IRX Therapeutics, Medpace, Merck-Serono, MSD, Novartis, Pfizer, Roche,  
635 and Buran; occasional fees for participation as a speaker at conferences/congresses or as a  
636 scientific consultant for advisory boards from AstraZeneca, Bayer, MSD, Merck-Serono,  
637 AccMed, Neutron Therapeutics, Inc., and Alentis. S.C. declares occasional fees for  
638 participation as a speaker at conferences/congresses from AccMed; support for attending  
639 meetings and/or travel from AccMed, MultiMed Engineers srl, Care Insight sas, unrelated to  
640 the content of this work. All remaining authors declare no competing interests.

641

## 642 References

- 643 1. Barsouk, A., Aluru, J. S., Rawla, P., Saginala, K. & Barsouk, A. Epidemiology, Risk Factors, and  
644 Prevention of Head and Neck Squamous Cell Carcinoma. *Med. Sci.* **11**, 42 (2023).
- 645 2. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality  
646 Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **71**, 209–249 (2021).
- 647 3. Gormley, M., Creaney, G., Schache, A., Ingarfield, K. & Conway, D. I. Reviewing the epidemiology  
648 of head and neck cancer: definitions, trends and risk factors. *Br. Dent. J.* **233**, 780–786 (2022).
- 649 4. Sabatini, M. E. & Chiocca, S. Human papillomavirus as a driver of head and neck cancers. *Br. J.*  
650 *Cancer* **122**, 306–314 (2020).
- 651 5. Leemans, C. R., Snijders, P. J. F. & Brakenhoff, R. H. The molecular landscape of head and neck  
652 cancer. *Nat. Rev. Cancer* **18**, 269–282 (2018).
- 653 6. Mes, S. W., Leemans, C. R. & Brakenhoff, R. H. Applications of molecular diagnostics for  
654 personalized treatment of head and neck cancer: state of the art. *Expert Rev. Mol. Diagn.* **16**,  
655 205–221 (2016).
- 656 7. Reid, P. *et al.* Diversity of cancer stem cells in head and neck carcinomas: The role of HPV in  
657 cancer stem cell heterogeneity, plasticity and treatment response. *Radiother. Oncol.* **135**, 1–12  
658 (2019).
- 659 8. Machiels, J.-P. *et al.* Squamous cell carcinoma of the oral cavity, larynx, oropharynx and  
660 hypopharynx: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and  
661 follow-up. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **31**, 1462–1475 (2020).
- 662 9. Gatta, G. *et al.* Burden and centralised treatment in Europe of rare tumours: results of  
663 RARECAREnet—a population-based study. *Lancet Oncol.* **18**, 1022–1039 (2017).
- 664 10. Ballman, K. V. Biomarker: Predictive or Prognostic? *J. Clin. Oncol.* **33**, 3968–3971 (2015).
- 665 11. Al-Tashi, Q. *et al.* Machine Learning Models for the Identification of Prognostic and Predictive  
666 Cancer Biomarkers: A Systematic Review. *Int. J. Mol. Sci.* **24**, 7781 (2023).



- 667 12. Huo, M. *et al.* Tumor microenvironment characterization in head and neck cancer identifies  
668 prognostic and immunotherapeutically relevant gene signatures. *Sci. Rep.* **10**, 11163 (2020).
- 669 13. Feng, B. & Hess, J. Immune-Related Mutational Landscape and Gene Signatures: Prognostic  
670 Value and Therapeutic Impact for Head and Neck Cancer. *Cancers* **13**, 1162 (2021).
- 671 14. Wang, J. *et al.* Six-gene signature for predicting survival in patients with head and neck  
672 squamous cell carcinoma. *Aging* **12**, 767–783 (2020).
- 673 15. Liu, B. *et al.* Prognostic Value of Eight-Gene Signature in Head and Neck Squamous Carcinoma.  
674 *Front. Oncol.* **11**, (2021).
- 675 16. Tonella, L., Giannoccaro, M., Alfieri, S., Canevari, S. & De Cecco, L. Gene Expression Signatures  
676 for Head and Neck Cancer Patient Stratification: Are Results Ready for Clinical Application? *Curr.*  
677 *Treat. Options Oncol.* **18**, 32 (2017).
- 678 17. Steckler, A. & McLeroy, K. R. The Importance of External Validity. *Am. J. Public Health* **98**, 9–10  
679 (2008).
- 680 18. De Cecco, L., Bossi, P., Locati, L., Canevari, S. & Licitra, L. Comprehensive gene expression meta-  
681 analysis of head and neck squamous cell carcinoma microarray data defines a robust survival  
682 predictor. *Ann. Oncol.* **25**, 1628–1635 (2014).
- 683 19. Cavalieri, S. *et al.* Clinical Validity of a Prognostic Gene Expression Cluster-Based Model in  
684 Human Papillomavirus–Positive Oropharyngeal Carcinoma. *JCO Precis. Oncol.* 1666–1676 (2021)  
685 doi:10.1200/PO.21.00094.
- 686 20. Locati *et al.* Mining of Self-Organizing Map Gene-Expression Portraits Reveals Prognostic  
687 Stratification of HPV-Positive Head and Neck Squamous Cell Carcinoma. *Cancers* **11**, 1057  
688 (2019).
- 689 21. Eschrich, S. A. *et al.* A Gene Expression Model of Intrinsic Tumor Radiosensitivity: Prediction of  
690 Response and Prognosis After Chemoradiation. *Int. J. Radiat. Oncol.* **75**, 489–496 (2009).
- 691 22. Torres-Roca, J. F., Grass, G. D., Scott, J. G. & Eschrich, S. A. Towards Data Driven RT Prescription:  
692 Integrating Genomics into RT Clinical Practice. *Semin. Radiat. Oncol.* **33**, 221–231 (2023).

- 693 23. Wells, J. D., Griffin, J. R. & Miller, T. W. Pan-Cancer Transcriptional Models Predicting  
694 Chemosensitivity in Human Tumors. *Cancer Inform.* **20**, 117693512110024 (2021).
- 695 24. De Cecco, L. *et al.* Head and neck cancer subtypes with biological and clinical relevance: Meta-  
696 analysis of gene-expression data. *Oncotarget* **6**, 9627–9642 (2015).
- 697 25. Carengo, A. *et al.* Gene Expression Clustering and Selected Head and Neck Cancer Gene  
698 Signatures Highlight Risk Probability Differences in Oral Premalignant Lesions. *Cells* **9**, 1828  
699 (2020).
- 700 26. Bossi, P. *et al.* Functional Genomics Uncover the Biology behind the Responsiveness of Head and  
701 Neck Squamous Cell Cancer Patients to Cetuximab. *Clin. Cancer Res.* **22**, 3961–3970 (2016).
- 702 27. Lenoci, D. *et al.* Biological properties of hypoxia-related gene expression models/signatures on  
703 clinical benefit of anti-EGFR treatment in two head and neck cancer window-of-opportunity  
704 trials. *Oral Oncol.* **126**, 105756 (2022).
- 705 28. Cavalieri, S. *et al.* Development of a multiomics database for personalized prognostic forecasting  
706 in head and neck cancer: The Big Data to Decide EU Project. *Head Neck* **43**, 601–612 (2021).
- 707 29. Sablin, M.-P. *et al.* Identification of new candidate therapeutic target genes in head and neck  
708 squamous cell carcinomas. *Oncotarget* **7**, 47418–47430 (2016).
- 709 30. Neuzillet, C. *et al.* Prognostic value of intratumoral *Fusobacterium nucleatum* and association  
710 with immune-related gene expression in oral squamous cell carcinoma patients. *Sci. Rep.* **11**,  
711 7870 (2021).
- 712 31. Hoffmann, C. *et al.* MMP2 as an independent prognostic stratifier in oral cavity cancers.  
713 *Oncoimmunology* **9**, 1754094 (2020).
- 714 32. Leblanc, O. *et al.* Biomarkers of cetuximab resistance in patients with head and neck squamous  
715 cell carcinoma. *Cancer Biol. Med.* **17**, 208–217 (2020).
- 716 33. Lecerf, C. *et al.* Immune gene expression in head and neck squamous cell carcinoma patients.  
717 *Eur. J. Cancer* **121**, 210–223 (2019).

- 718 34. Hoffmann, C. *et al.* PD-L1 and ICOSL discriminate human Secretary and Helper dendritic cells in  
719 cancer, allergy and autoimmunity. *Nat. Commun.* **13**, 1983 (2022).
- 720 35. Moreira, A. *et al.* Prognostic value of tumor mutational burden in patients with oral cavity  
721 squamous cell carcinoma treated with upfront surgery. *ESMO Open* **6**, 100178 (2021).
- 722 36. Karabajakian, A. *et al.* Longitudinal assessment of PD-L1 expression and gene expression profiles  
723 in patients with head and neck cancer reveals temporal heterogeneity. *Oral Oncol.* **119**, 105368  
724 (2021).
- 725 37. The Cancer Genome Atlas (TCGA). <https://www.cancer.gov/tcga>.
- 726 38. Lawrence, M. S. *et al.* Comprehensive genomic characterization of head and neck squamous cell  
727 carcinomas. *Nature* **517**, 576–582 (2015).
- 728 39. Lohavanichbutr, P. *et al.* A 13-Gene Signature Prognostic of HPV-Negative OSCC: Discovery and  
729 External Validation. *Clin. Cancer Res.* **19**, 1197–1203 (2013).
- 730 40. Brazma, A. *et al.* Minimum information about a microarray experiment (MIAME)—toward  
731 standards for microarray data. *Nat. Genet.* **29**, 365–371 (2001).
- 732 41. Brazma, A. *et al.* MINSEQE: Minimum Information about a high-throughput Nucleotide  
733 SeQuencing Experiment - a proposal for standards in functional genomic data reporting. (2012)  
734 doi:10.5281/zenodo.5706412.
- 735 42. Edge, S. B. & Compton, C. C. The American Joint Committee on Cancer: the 7th Edition of the  
736 AJCC Cancer Staging Manual and the Future of TNM. *Ann. Surg. Oncol.* **17**, 1471–1474 (2010).
- 737 43. Hernández, L. *et al.* HeNeCO: An ontology for integrative research in Head and Neck cancer.  
738 *Int. J. Med. Inf.* **181**, 105284 (2024).
- 739 44. Nauta, I. H. *et al.* The unveiled reality of human papillomavirus as risk factor for oral cavity  
740 squamous cell carcinoma. *Int. J. Cancer* **149**, 420–430 (2021).
- 741 45. Jones, M. & Fowler, R. Immortal time bias in observational studies of time-to-event outcomes. *J.*  
742 *Crit. Care* **36**, 195–199 (2016).

- 743 46. Ness, A. R. *et al.* Establishing a large prospective clinical cohort in people with head and neck  
744 cancer as a biomedical resource: head and neck 5000. *BMC Cancer* **14**, 973 (2014).
- 745 47. Ness, A. R. *et al.* Recruitment, response rates and characteristics of 5511 people enrolled in a  
746 prospective clinical cohort study: head and neck 5000. *Clin. Otolaryngol.* **41**, 804–809 (2016).
- 747 48. Mi, X., Hammill, B. G., Curtis, L. H., Lai, E. C.-C. & Setoguchi, S. Use of the landmark method to  
748 address immortal person-time bias in comparative effectiveness research: a simulation study.  
749 *Stat. Med.* **35**, 4824–4836 (2016).
- 750 49. Gleiss, A., Oberbauer, R. & Heinze, G. An unjustified benefit: immortal time bias in the analysis  
751 of time-dependent events. *Transpl. Int.* **31**, 125–130 (2018).
- 752 50. Lash, T. L., Fink, A. K. & Fox, M. P. Misclassification. in *Applying Quantitative Bias Analysis to*  
753 *Epidemiologic Data* (eds. Lash, T. L., Fox, M. P. & Fink, A. K.) 79–108 (Springer, New York, NY,  
754 2009). doi:10.1007/978-0-387-87959-8\_6.
- 755 51. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and  
756 microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
- 757 52. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis.  
758 *BMC Bioinformatics* **9**, 559 (2008).
- 759 53. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using  
760 empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
- 761 54. Zhang, Y., Jenkins, D. F., Manimaran, S. & Johnson, W. E. Alternative empirical Bayes models for  
762 adjusting for batch effects in genomic studies. *BMC Bioinformatics* **19**, 262 (2018).
- 763 55. Nygaard, V., Rødland, E. A. & Hovig, E. Methods that remove batch effects while retaining group  
764 differences may lead to exaggerated confidence in downstream analyses. *Biostatistics* **17**, 29–39  
765 (2016).
- 766 56. Carengo, A. *et al.* hacksig: a unified and tidy R framework to easily compute gene expression  
767 signature scores. *Bioinformatics* **38**, 2940–2942 (2022).

- 768 57. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for  
769 Statistical Computing, Vienna, Austria, 2022).
- 770 58. Mitra, R. *et al.* Learning from data with structured missingness. *Nat. Mach. Intell.* **5**, 13–23  
771 (2023).
- 772 59. Harrell, F. E., Jr, Califf, R. M., Pryor, D. B., Lee, K. L. & Rosati, R. A. Evaluating the Yield of Medical  
773 Tests. *JAMA* **247**, 2543–2546 (1982).
- 774 60. Hartman, N., Kim, S., He, K. & Kalbfleisch, J. D. Pitfalls of the concordance index for survival  
775 outcomes. *Stat. Med.* **42**, 2179–2190 (2023).
- 776 61. Royston, P. Explained Variation for Survival Models. *Stata J.* **6**, 83–96 (2006).
- 777 62. O’Quigley, J., Xu, R. & Stare, J. Explained randomness in proportional hazards models. *Stat. Med.*  
778 **24**, 479–489 (2005).
- 779 63. Nagelkerke, N. J. D. A note on a general definition of the coefficient of determination.  
780 *Biometrika* **78**, 691–692 (1991).
- 781 64. Shuster, J. J. Median follow-up in clinical trials. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **9**, 191–  
782 192 (1991).
- 783 65. Dai, Y.-H. *et al.* Radiosensitivity index emerges as a potential biomarker for combined  
784 radiotherapy and immunotherapy. *Npj Genomic Med.* **6**, 40 (2021).
- 785 66. Torres-Roca, J. F. A molecular assay of tumor radiosensitivity: a roadmap towards biology-based  
786 personalized radiation therapy. *Pers. Med.* **9**, 547–557 (2012).
- 787 67. Groenwold, R. H. H. *et al.* Missing covariate data in clinical research: when and when not to use  
788 the missing-indicator method for analysis. *CMAJ* **184**, 1265–1269 (2012).
- 789 68. Song, M., Zhou, X., Pazaris, M. & Spiegelman, D. The Missing Covariate Indicator Method is  
790 Nearly Valid Almost Always. Preprint at <https://doi.org/10.48550/arXiv.2111.00138> (2021).
- 791 69. Merck. Merck Provides Update on Xevinapant Program in Locally Advanced Head and Neck  
792 Cancer [Press release]. <https://www.merckgroup.com/en/news/xevinapant-update.html> (2024).

- 793 70. Lee, N. Y. *et al.* Avelumab plus standard-of-care chemoradiotherapy versus chemoradiotherapy  
794 alone in patients with locally advanced squamous cell carcinoma of the head and neck: a  
795 randomised, double-blind, placebo-controlled, multicentre, phase 3 trial. *Lancet Oncol.* **22**, 450–  
796 462 (2021).
- 797 71. Machiels, J.-P. *et al.* Pembrolizumab plus concurrent chemoradiotherapy versus placebo plus  
798 concurrent chemoradiotherapy in patients with locally advanced squamous cell carcinoma of  
799 the head and neck (KEYNOTE-412): a randomised, double-blind, phase 3 trial. *Lancet Oncol.* **25**,  
800 572–587 (2024).
- 801 72. Mehanna, H. *et al.* Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-  
802 positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase  
803 3 trial. *The Lancet* **393**, 51–60 (2019).
- 804 73. Gillison, M. L. *et al.* Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive  
805 oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority  
806 trial. *The Lancet* **393**, 40–50 (2019).
- 807 74. Tao, Y. *et al.* Pembrolizumab versus cetuximab concurrent with radiotherapy in patients with  
808 locally advanced squamous cell carcinoma of head and neck unfit for cisplatin (GORTEC 2015-01  
809 PembroRad): a multicenter, randomized, phase II trial. *Ann. Oncol.* **34**, 101–110 (2023).
- 810  
811

812 **Table 1.** Cohort characteristics for head and neck cancer patients used to test the *172-GS* and *3*  
813 *clusters HPV* signatures. The disease-free survival (DFS) dataset is a subset of the overall survival  
814 dataset, where patients without data on DFS are excluded.

	Overall survival dataset (N=1097)	Disease-free survival dataset (N=907)
<b>172-GS score</b>		
Mean (SD)	0.0 (1.0)	0.0 (1.0)
Median (Min, Max)	0.1 (-4.5, 3.6)	0.1 (-4.9, 3.5)
<b>3 clusters HPV score</b>		
Mean (SD)	0.0 (1.0)	0.0 (1.0)
Median (Min, Max)	0.0 (-2.8, 6.9)	0.0 (-2.9, 3.5)
<b>Study ID</b>		
BD2_INT_MI	310 (28.3%)	310 (34.2%)
BD2_UDUS	127 (11.6%)	127 (14.0%)
GECO110	107 (9.8%)	107 (11.8%)
GSE41613	97 (8.8%)	0 (0%)
SCANDARE	32 (2.9%)	0 (0%)
TCGA	424 (38.7%)	363 (40.0%)
<b>Clinical age at diagnosis</b>		
Mean (SD)	61.0 (11.9)	60.8 (11.7)
Median (Min, Max)	61.0 (19.0, 93.0)	61.0 (19.0, 93.0)
<b>Clinical sex</b>		
Male	794 (72.4%)	660 (72.8%)
Female	303 (27.6%)	247 (27.2%)
<b>Disease extension at diagnosis</b>		
Early disease	161 (14.7%)	98 (10.8%)
Locoregionally advanced disease	936 (85.3%)	809 (89.2%)
<b>Tumor region</b>		
Oropharynx	227 (20.7%)	218 (24.0%)
Hypopharynx	55 (5.0%)	52 (5.7%)
Larynx	165 (15.0%)	150 (16.5%)
Oral cavity	598 (54.5%)	451 (49.7%)
Missing	52 (4.7%)	36 (4.0%)
<b>HPV status</b>		
Positive	156 (14.2%)	151 (16.6%)
Negative	941 (85.8%)	756 (83.4%)
<b>Smoking</b>		
Never	230 (21.0%)	213 (23.5%)
Current or Former	653 (59.5%)	609 (67.1%)
Missing	214 (19.5%)	85 (9.4%)
<b>Undergone cancer surgery</b>		
No	187 (17.0%)	187 (20.6%)
Yes	854 (77.8%)	672 (74.1%)
Missing	56 (5.1%)	48 (5.3%)
<b>Radiotherapy treatment</b>		
No	208 (19.0%)	146 (16.1%)
Yes	752 (68.6%)	633 (69.8%)
Missing	137 (12.5%)	128 (14.1%)
<b>Systemic treatment</b>		
No	302 (27.5%)	233 (25.7%)
Yes	461 (42.0%)	426 (47.0%)
Missing	334 (30.4%)	248 (27.3%)

815

816

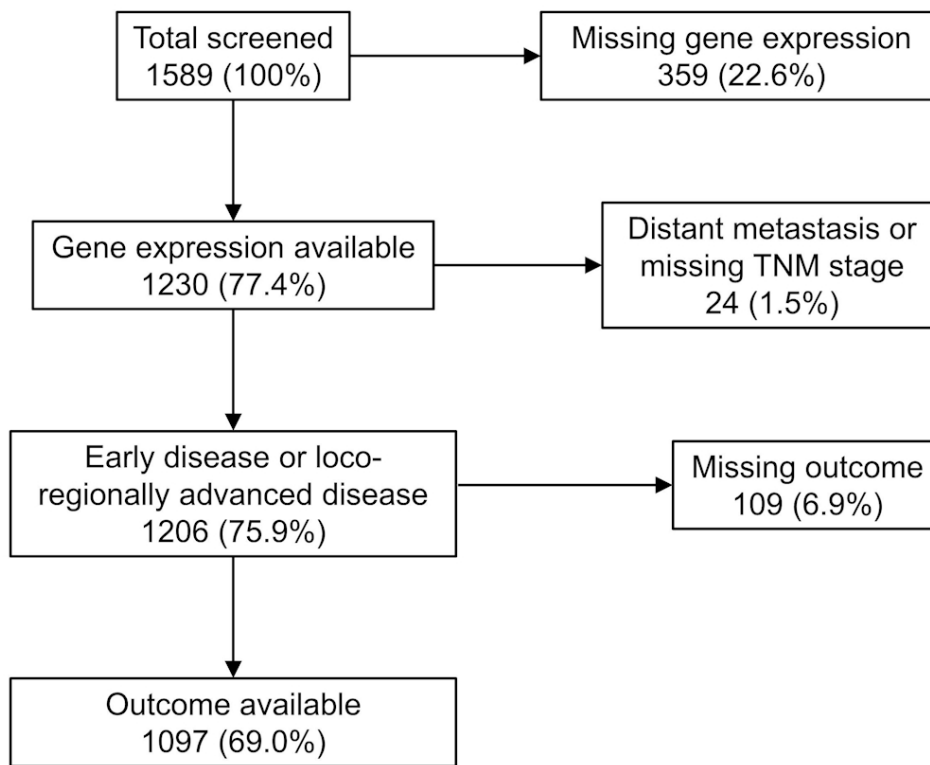
817 **Table 2.** Hazard ratios (HR) of gene signatures and covariates that were tested in interaction  
818 with the signatures for head and neck cancer patients. Estimates are from models where 2-  
819 year overall survival or 2-year disease-free survival was the endpoint and the effect of other  
820 clinical covariates were jointly estimated (see Tables S5, S6, S7, S8, S9 for detailed results).

Signature	Variable	Parameter	Overall survival		Disease-free survival	
			HR [95% CI]	P-value	HR [95% CI]	P-value
<i>172-GS</i>	HPV-status	HPV positive	Reference	-	Reference	-
		HPV negative	4.27 [2.03, 9.01]	< <b>0.001</b>	4.43 [2.44, 8.02]	< <b>0.001</b>
	Signature score	Score given HPV positive	1.87 [0.61, 5.78]	0.274	1.67 [0.69, 4.05]	0.254
		Score given HPV negative	1.41 [1.25, 1.59]	< <b>0.001</b>	1.36 [1.17, 1.59]	< <b>0.001</b>
<i>3 clusters HPV</i>	HPV-status	HPV positive	Reference	-	Reference	-
		HPV negative	4.17 [2.03, 8.57]	< <b>0.001</b>	4.01 [2.26, 7.14]	< <b>0.001</b>
	Signature score	Score given HPV positive	0.41 [0.24, 0.71]	<b>0.001</b>	0.51 [0.33, 0.77]	<b>0.001</b>
		Score given HPV negative	0.89 [0.77, 1.02]	0.084	1.06 [0.91, 1.23]	0.454
<i>RSI</i>	Radiotherapy	No	Reference	-	Reference	-
		Yes	0.81 [0.54, 1.21]	0.306	1.08 [0.82, 1.42]	0.601
	Signature score	Score given radiotherapy = no	1.17 [0.90, 1.52]	0.249	1.01 [0.79, 1.27]	0.929
		Score given radiotherapy = yes	1.25 [1.08, 1.44]	<b>0.003</b>	1.08 [0.82, 1.42]	0.306
<i>Pancancer-cisplatin</i>	Systemic therapy	Platinum based	Reference	-	Reference	-
		Non-platinum based	0.57 [0.28, 1.15]	0.114	0.94 [0.44, 2.01]	0.873
		No systemic treatment	1.24 [0.75, 2.04]	0.409	0.97 [0.63, 1.49]	0.882
	Signature score	Score given systemic therapy = platinum based	1.29 [1.04, 1.59]	<b>0.020</b>	1.23 [0.82, 1.84]	0.316
		Score given systemic therapy = non-platinum based	0.82 [0.42, 1.59]	0.552	2.00 [0.90, 4.45]	0.089
		Score given systemic therapy = no	1.37 [0.51, 3.69]	0.540	0.92 [0.37, 2.31]	0.862
<i>CI3-hypoxia</i>	Systemic therapy	Cetuximab based	Reference	-	Reference	-
		Non-cetuximab based	1.07 [0.49, 2.31]	0.869	0.71 [0.34, 1.48]	0.363
		No systemic treatment	1.22 [0.51, 2.91]	0.657	0.69 [0.31, 1.54]	0.364
	Signature score	Score given systemic therapy = cetuximab based	0.67 [0.39, 1.14]	0.137	0.61 [0.33, 1.13]	0.114
		Score given systemic therapy = non-cetuximab based	0.69 [0.54, 0.87]	<b>0.002</b>	0.59 [0.44, 0.80]	<b>0.001</b>
		Score given systemic therapy = no	1.70 [1.12, 2.58]	<b>0.013</b>	1.71 [1.14, 2.56]	<b>0.010</b>

821



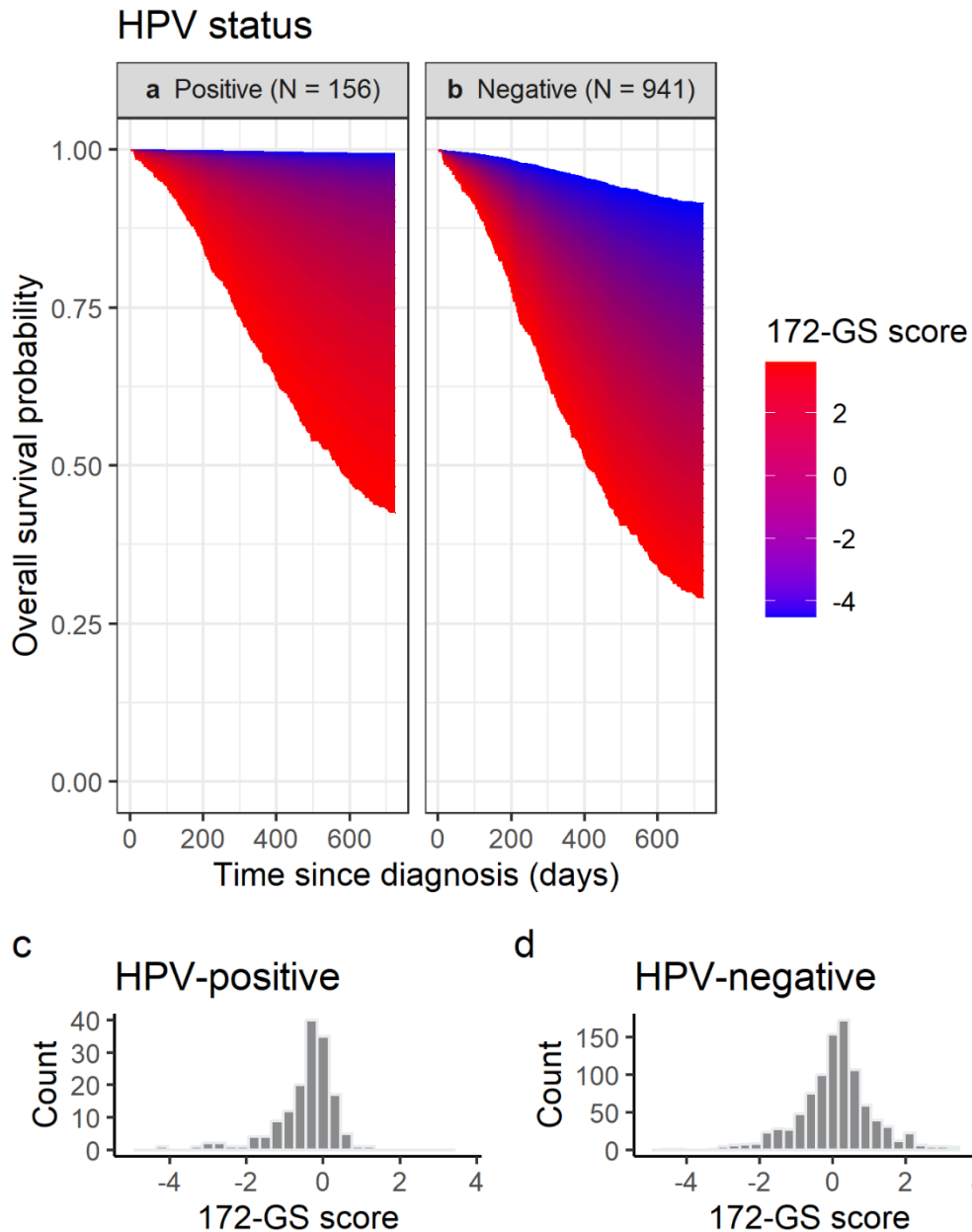
822  
823



824

825 **Figure 1. Exclusion flowchart.** The largest number of eligible patients used in the most  
826 inclusive models (models testing the signatures *172-GS*, *3 clusters HPV* and *RSI* with overall  
827 survival as the endpoint) is shown. See Figures S1, S2, S3 for exclusion flowcharts for other  
828 gene signature models.

829



830

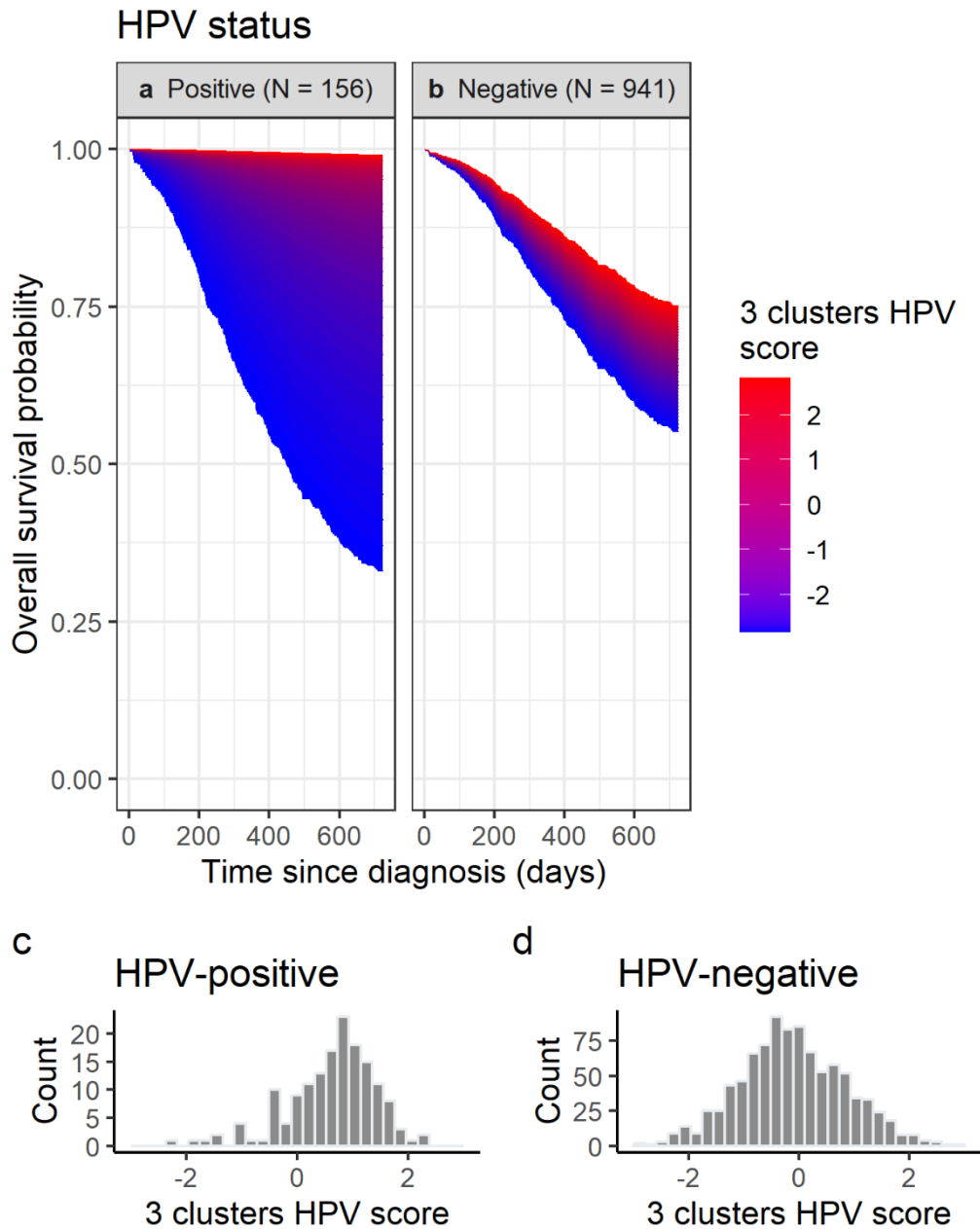
831 **Figure 2. Estimated relationship between the 172-GS signature and 2-year overall**  
832 **survival. a, b** Estimated relationship in HPV-positive patients and HPV-negative patients,  
833 respectively. The signature score is scaled to a mean of 0 and a standard deviation of 1. Note  
834 the difference in sample size (N) per group. **c, d** Distribution of signature values in HPV-  
835 positive and HPV-negative patients, respectively.

836

837

838

839

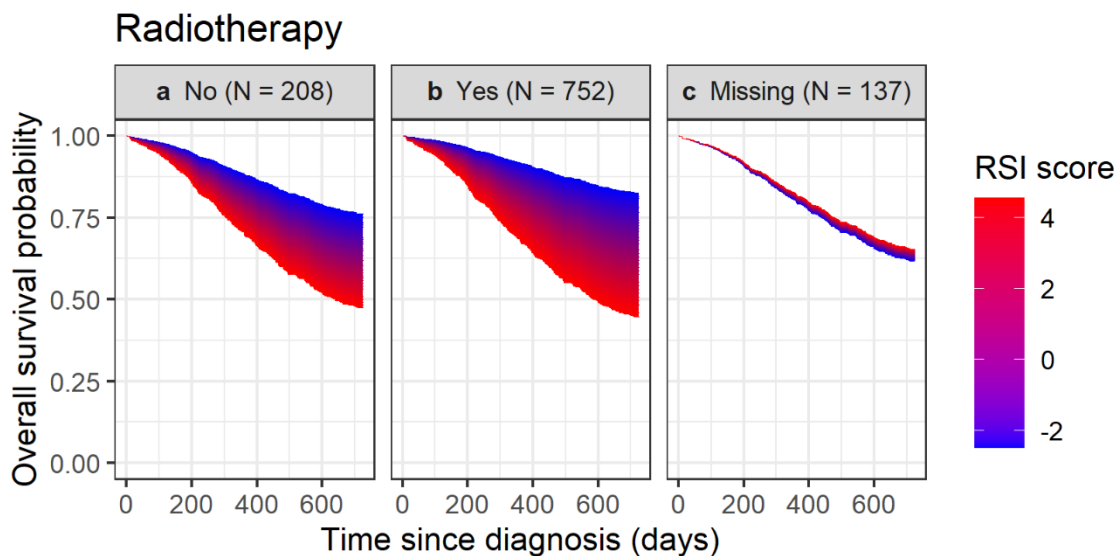


840

841 **Figure 3. Estimated relationship between the 3 clusters HPV signature and 2-year**  
842 **overall survival. a, b** Estimated relationship in HPV-positive patients and HPV-negative  
843 patients, respectively. The signature score is scaled to a mean of 0 and a standard deviation of  
844 1. Note the difference in sample size (N) per group. **c, d** Distribution of signature values in  
845 HPV-positive and HPV-negative patients, respectively.

846

847  
848

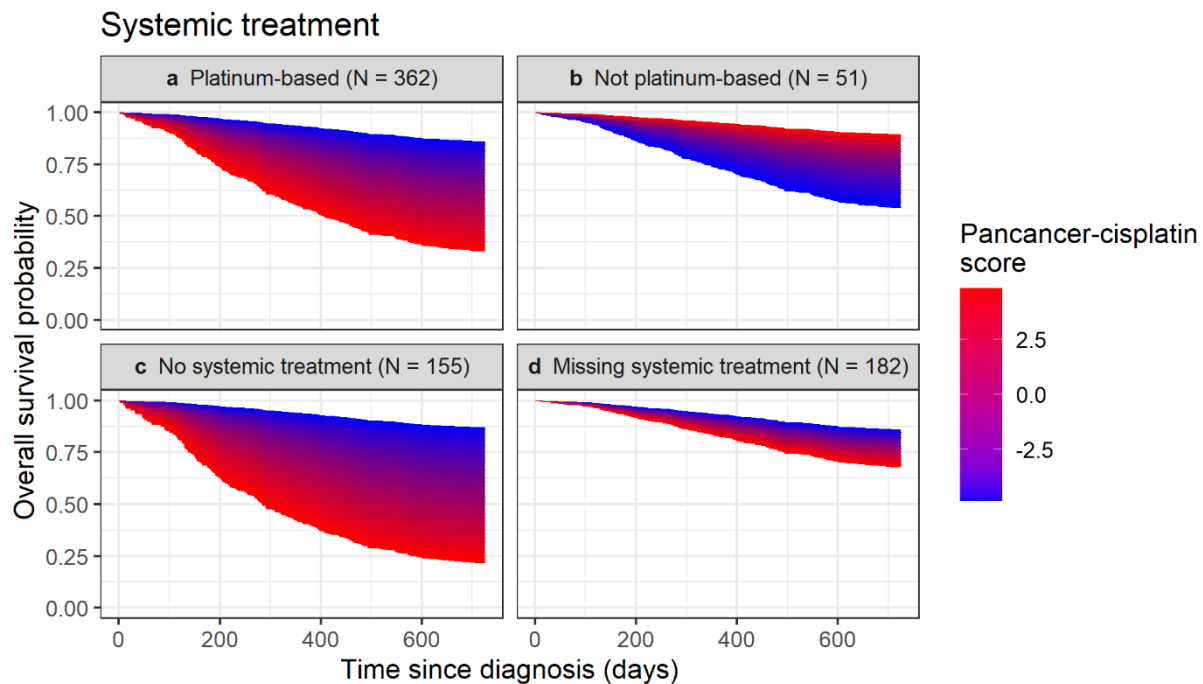


849

850 **Figure 4. Estimated relationship between the radiosensitivity index (*RSI*) and 2-year**  
851 **overall survival. a** Relationship in patients that did not receive radiotherapy. **b** Relationship  
852 in patients that received radiotherapy. **c** Relationship in patients with missing information  
853 about radiotherapy treatment. The *RSI* score is scaled to a mean of 0 and a standard deviation  
854 of 1. Note the difference in sample size (*N*) per group.

855

856



857

858

859

860

861

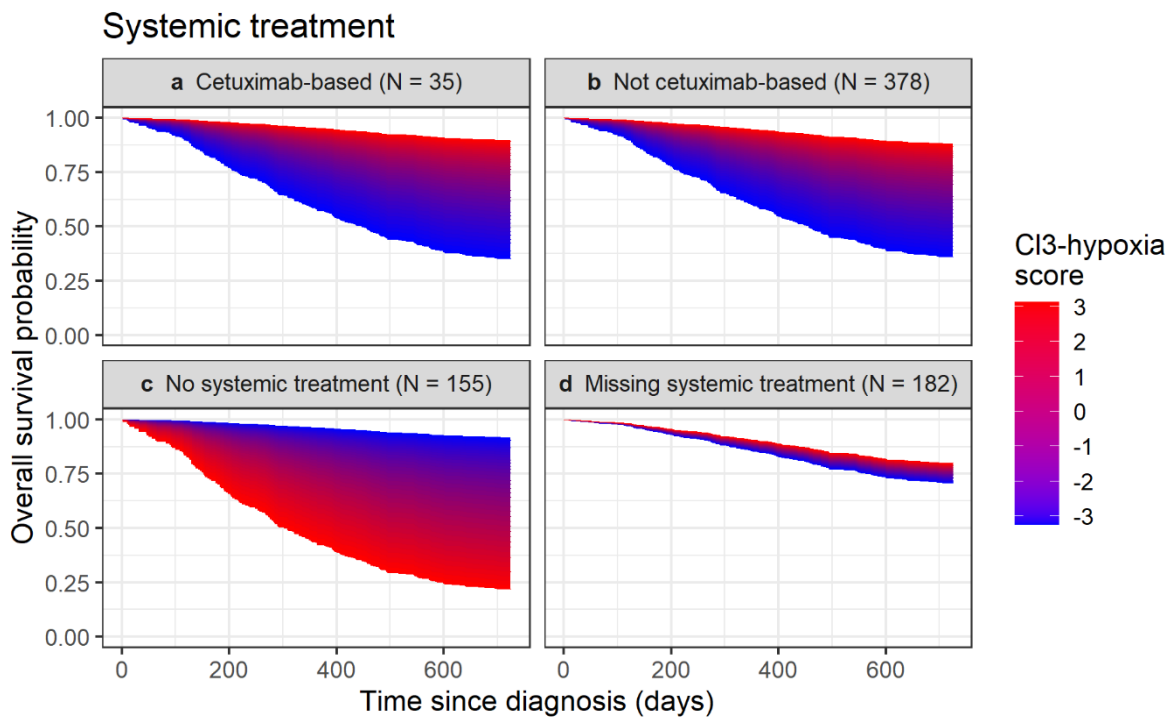
862

863

864

**Figure 5. Estimated relationship between the *pancancer-cisplatin* signature and 2-year overall survival. a** Relationship in patients that received platinum-based chemotherapy. **b** Relationship in patients that received non-platinum based systemic treatment. **c** Relationship in patients that did not receive systemic treatment. **d** Relationship in patients with missing information about systemic treatment. The signature score is scaled to a mean of 0 and a standard deviation of 1. Note the difference in sample size (N) per group.

865



866

867

868

869

870

871

872

**Figure 6. Estimated relationship between the *CI3-hypoxia* signature and 2-year overall survival.** **a** Relationship in patients that received cetuximab-based treatment. **b** Relationship in patients that received non-cetuximab based systemic treatment. **c** Relationship in patients that did not receive systemic treatment. **d** Relationship in patients with missing information about systemic treatment. The signature score is scaled to a mean of 0 and a standard deviation of 1. Note the difference in sample size (N) per group.

873