PREDICTING THE CLINICAL OUTCOME OF ORAL POTENTIALLY MALIGNANT DISORDERS USING TRANSCRIPTOMIC-BASED MOLECULAR PATHOLOGY

Hans Prakash Sathasivam[1,2,*], Ralf Kist[1,3,*], Philip Sloan[1,4], Peter Thomson[5], Michael Nugent[6], John Alexander[7], Syed Haider[7,*], Max Robinson[1,4,*]

[1] School of Dental Sciences, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.
[2] Cancer Research Centre, Institute for Medical Research, National Institute of Health, Setia Alam, Malaysia.
[3] Newcastle University Biosciences Institute, Newcastle University Centre for Cancer, Newcastle upon Tyne, UK.
[4] Department of Cellular Pathology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK.
[5] Oral and Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong, Hong Kong SAR.
[6] Oral and Maxillofacial Surgery, Sunderland Royal Hospital, Sunderland, UK.
[7] The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London SW3 6JB, UK.

*These authors contributed equally to this work.

Corresponding author
Dr Max Robinson
Department of Cellular Pathology
Newcastle upon Tyne Hospitals NHS Foundation Trust
Royal Victoria Infirmary
Newcastle upon Tyne
NE1 4LP
Tel: +44 0191 282 9807
Email: max.robinson@nhs.net
Orcid: 0000-0003-4491-6865

Conflicts of interest

35 The authors have no conflicts of interest to declare.

36

37 **Abstract**

38 **Background:** This study was undertaken to develop and validate a gene expression
39 signature that characterises oral potentially malignant disorders (OPMD) with a high
40 risk of undergoing malignant transformation.

41 **Methods:** Patients with oral epithelial dysplasia at one hospital were selected as the
42 'training set' (n=56) whilst those at another hospital were selected for the 'test set'
43 (n=66). RNA was extracted from formalin-fixed paraffin-embedded (FFPE) diagnostic
44 biopsies and analysed using the NanoString nCounter platform. A targeted panel of
45 42 genes selected on their association with oral carcinogenesis was used to develop
46 a prognostic gene signature. Following data normalisation, uni- and multivariable
47 analysis, as well as prognostic modelling, were employed to develop and validate the
48 gene signature.

49 **Results:** A prognostic classifier composed of 11 genes was developed using the
50 training set. The multivariable prognostic model was used to predict patient risk
51 scores in the test set. The prognostic gene signature was an independent predictor
52 of malignant transformation when assessed in the test set, with the high-risk group
53 showing worse prognosis [Hazard ratio = 12.65, p = 0.0003].

54 **Conclusions:** This study demonstrates proof of principle that RNA extracted from
55 FFPE diagnostic biopsies of OPMD, when analysed on the NanoString nCounter
56 platform, can be used to generate a molecular classifier that stratifies the risk of
57 malignant transformation with promising clinical utility.

58

59 **Keywords**

60 Oral potentially malignant disorders: oral epithelial dysplasia: gene expression: gene
61 signature: prognostic modelling: Nanostring.

62

63 **Background**

64 Oral squamous cell carcinoma (OSCC) has a high rate of morbidity and mortality
65 worldwide.[1-5] Around 30-50% of patients with OSCC die from the disease within five
66 years and survival rates have not improved over many decades.[2, 5] Such adverse
67 outcomes have mostly been attributed to late presentation of the disease, as early
68 stage disease can be cured with effective treatment.[1, 2, 6] Early detection of OSCC is

69 feasible as they are usually preceded by clinically identifiable lesions termed ‘oral
70 potentially malignant disorders’ (OPMD).[1, 2, 7, 8]
71 OPMD are defined as clinical disorders having an increased risk of developing
72 OSCC compared to clinically ‘normal’ oral mucosa.[1, 7] The majority of OPMD do not
73 transform to cancer, consequently the challenge is identifying those lesions that are
74 most likely to undergo malignant transformation.[9-12]
75 Clinical and histopathological features, though informative, are not very accurate in
76 predicting the clinical behaviour of these lesions.[13] Nevertheless, currently the
77 presence and grade of OED is considered to be the most useful indicator of
78 malignant transformation in OPMD and provides the basis for patient stratification
79 endorsed by the World Health Organisation.[1] A systematic review and meta-analysis
80 indicates that excision of oral dysplastic lesions reduces the risk of malignant
81 transformation by approximately 3-fold.[11] Generally severe epithelial dysplasia or
82 high-grade epithelial dysplasia is treated empirically by surgical excision,[14-16]
83 however, it is not clear how patient outcomes can be improved across all grades of
84 dysplasia and those patients with non-dysplastic OPMD. Currently, it is unknown
85 whether all OPMD should be excised or if only certain lesions benefit from a surgical
86 intervention.
87 Numerous studies have assessed the prognostic ability of various biomarkers in
88 OPMD, however, no molecular test has proved to be particularly useful in clinical
89 practice.[17-23] Discovering a molecular signature that is altered in OPMD and
90 indicative of the progression to oral cancer could facilitate personalized
91 management protocols for individual patients.
92 Contemporary gene expression profiling is being used to develop prognostic and
93 predictive gene signatures in various cancers, including head and neck cancers.[24, 25]
94 A study by Saintigny et al (2011) proposed a gene expression-based prediction
95 model for OPMD that showed superior prognostic accuracy when compared to
96 models using clinico-pathological risk factors alone.[26] However, the patients in their
97 study were enrolled in a clinical trial in which some patients received active
98 intervention in the form of drugs that may have influenced clinical outcome and gene
99 expression.[26] Furthermore, the findings of their study have yet to be validated.
100 Whilst formalin-fixed paraffin-embedded (FFPE) tissue is an invaluable resource
101 linked to longitudinal disease-related outcome; it is often not possible to extract
102 adequate amounts of high-quality nucleic acid for downstream analysis. A novel

103 gene expression profiling system that relies on direct measurement of transcripts
104 using colour-coded oligonucleotide probes producing molecular barcodes, the
105 NanoString nCounter platform (NanoString Technologies, Seattle, USA), has been
106 able to provide accurate gene expression data using RNA obtained from FFPE
107 material.[27, 28] Recent studies have shown that mRNA expression analysis using the
108 NanoString platform were equivalent to that achieved through quantitative real-time
109 polymerase chain reaction (qPCR) and possibly superior to microarrays.[27-31]
110 Furthermore, the Prosigna™ breast cancer prognostic gene signature assay is
111 based on Nanostring technology and is approved by the US Food and Drug
112 Administration and recommended by UK National Institute for Health and Care
113 Excellence. The test is used to guide adjuvant chemotherapy decisions for women
114 with oestrogen receptor-positive, human epidermal growth factor receptor 2-negative
115 and lymph node-negative early breast cancer.
116 Despite the global health burden and relatively poor prognosis associated with
117 OSCC, a robust prognostic biomarker or prognostic model for predicting malignant
118 transformation in OPMD has yet to be developed and validated. This study was
119 undertaken to discover and then validate a transcriptomic-signature that identifies
120 OPMD with a high-risk of undergoing malignant transformation using FFPE-derived
121 RNA analysed on the NanoString nCounter platform.
122
123 **Methods**
124 ***Inclusion and exclusion criteria***
125 Consecutive OPMD cases were identified from a database at Newcastle University.
126 Cases with any one of the following characteristics were excluded: i) patients with
127 hereditary conditions that are linked to an increased risk of head and neck SCC
128 (such as ataxia telangiectasia, xeroderma pigmentosum, Fanconi anaemia); ii)
129 history of head and neck cancer; iii) history of radiotherapy to the head and neck
130 region; iv) patients that were diagnosed as having chronic hyperplastic
131 candidosis/chronic candidosis.
132 OPMD were classified as having undergone malignant transformation (MT) when
133 there was progression from an OPMD to oral squamous cell carcinoma (OSCC) after
134 a period of six months or more from the time of initial diagnosis. Those patients with
135 OPMD who were recorded as not having developed OSCC at their last known follow-
136 up appointment were classified as non-transforming (NT) cases with the caveat that

137 the patients were followed up for at least 12 months after diagnosis. All cases were
138 assessed for high-risk human papillomavirus (HR-HPV), and positive cases were
139 excluded from the study.

140 ***Patients***

141 Patients were selected from a database containing patients from two different
142 hospitals: i) Newcastle upon Tyne Hospitals NHS Foundation Trust and; ii) City
143 Hospitals Sunderland NHS Foundation Trust. Patients from Newcastle Hospitals
144 were selected as the 'training set' whilst patients from Sunderland Hospitals were
145 selected for the 'test set'.

146 ***Clinico-pathological data***

147 Demographic and clinico-pathological features as well as outcome data were
148 recorded for all cases. The following data points were collected and entered into a
149 Microsoft Excel spreadsheet: i) age at first diagnosis of OPMD: ii) sex: iii) clinical
150 diagnosis of lesion: iv) clinical outcome of OPMD: v) date of malignant
151 transformation or last follow-up: vi) World Health Organization (WHO) 2017 oral
152 epithelial dysplasia (OED) grading: vii) binary OED grading.

153 OED grading was performed following a modified three-tier system adapted from the
154 work published by Speight *et. al.* (2015).[32] The cases were graded using two different
155 classification systems: i) WHO 2017 (mild, moderate or severe):[1] binary (low-grade
156 or high-grade).[1, 33] All data were coded, link-anonymised and stored in password
157 protected computer files.

158 ***RNA extraction***

159 10 µm sections were cut from the FFPE blocks and placed in 2 ml microcentrifuge
160 tubes after discarding the first two sections. Whole sections that included both
161 epithelium and underlying connective tissue were used. The number of sections per
162 sample was dependent on the size of the FFPE tissue; as a guide four sections were
163 taken for small samples (<5 mm of epithelium), three for medium samples (5-10 mm)
164 and two for larger samples (>10 mm). If the amount of RNA extracted was not
165 sufficient, RNA extraction was repeated using an increased number of sections. RNA
166 extraction was performed using the RNeasy® FFPE kit (QIAGEN, Manchester, UK)
167 according to the manufacturer's protocol. FFPE blocks were sectioned immediately
168 before RNA extraction. The concentration and the quality of the isolated RNA were
169 measured using a NanoDrop 2000 Spectrophotometer (Thermo Fisher
170 Scientific,Swindon, UK). RNA was diluted to 150 ng/µL, aliquoted and stored in a -

171 80°C freezer prior to NanoString assay. RNA with a 260/280 ratio of 1.7 - 2.3 as well
172 as a 260/230 ratio in the range of 1.8 - 2.3 were considered to be of acceptable
173 quality for downstream assays.[34] RNA content for all samples was normalised to 30
174 ng/µl, and 150 ng of total RNA per sample was used for the assay.

175 ***NanoString nCounter Customised Panel***

176 A list of target genes for the NanoString nCounter Customised Panel (42 genes; 38
177 target and 4 housekeeping genes) was compiled based on the results of previous
178 experiments: a whole-transcriptome analysis with total RNA sequencing (RNA-
179 Seq),[35] results of previous differential gene expression work performed by our group
180 and review of published literature. The selection of candidate genes was discussed
181 and finalised through consensus by the authors; the gene list is shown in
182 Supplementary Table 1. Housekeeping / internal reference genes were selected on
183 the basis of low variation and even coverage across samples.[36-39]

184 ***NanoString nCounter hybridisation***

185 The NanoString nCounter platform uses hybridisation of short length probes (35- to
186 50- base sequence) that are subsequently fixed to a biotin-coated cartridge which is
187 then digitally imaged and counted to quantify mRNA expression. In-depth details
188 regarding NanoString technology can be obtained from Geiss et. al. (2008).[27]
189 NanoString assay was carried out at the Newcastle NanoString Unit, Newcastle
190 University using the nCounter MAX/FLEX system (NanoString Technologies, Seattle,
191 Washington, USA). Each assay comes with engineered External RNA Controls
192 Consortium (ERCC) synthetic internal negative and positive control probes. The
193 summarised laboratory workflow for the Customised CodeSet Panel gene
194 expression assay according to the manufacturer's protocol is outlined in the
195 Supplementary Methods.[34]

196 ***Normalisation of data and development of prognostic gene signature***

197 Nanostring profiling of codeset was pre-processed using R package
198 NanoStringNorm version 1.2.1. Data were assessed for batch effects using R
199 package FactoMineR version 1.39. Data were normalised using grid search over
200 parameter space as detailed previously,[40] resulting in the choice of parameters:
201 ‘geometric mean’ of positive controls, ‘mean’ of negative controls and ‘geometric
202 mean’ of top genes, and finally log2 transformed. Genes with zero counts in > 50%
203 samples were removed from subsequent analyses. This resulted in the removal of

204 genes: *CDKN2A, MMP1, DSPP, CERS1* and *IBSP*. All visualisations were generated
205 in R statistical environment version 3.6.1.

206 ***Statistical analysis and multivariable prognostic/survival modelling***

207 Statistical analysis and prognostic model building were performed using IBM SPSS
208 for Windows (version 24, IBM-SPSS Inc., Chicago, Illinois, USA) and the R statistical
209 environment version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).
210 Continuous data were always assessed for normality of distribution prior to choosing
211 appropriate statistical tests. Parametric and non-parametric tests were used for initial
212 analysis of demographic, clinical, pathological and molecular variables. For
213 continuous data, descriptive results were appropriately expressed as either median
214 with interquartile range (IQR) or mean with standard deviation (SD). For cross
215 tabulations and chi-squared tests, exact p-values were calculated where possible.
216 The Newcastle cohort was used as training set, while the Sunderland cohort was
217 used as a held-out test set. mRNA abundance data for genes was transformed to z-
218 scores. A multivariable generalised linear model with *L1*-penalty was fitted in cross-
219 validation (four-fold) settings to identify features predictive of patient outcome. This
220 process was repeated 100 times to select optimal *lambda* minimising cross-
221 validation error. The final model was used to predict risk scores in the test set and
222 predicted risk scores were dichotomised into risk groups (using median risk score
223 from the training set). These risk groups were tested for association with patient
224 outcome using Cox proportional hazards model. Survival modelling was performed
225 using R packages survival version 3.1-12 and glmnet version 4.0. All visualisations
226 were generated in R statistical environment version 3.6.1. Survival model was
227 adjusted for age, sex, site, OED grade and type of OPMD.
228 All statistical tests were two-sided, and results were considered statistically
229 significant at $p < 0.05$ value unless stated otherwise.
230

231 **Results**

232 Of the cases that fulfilled the selection criteria, 134 cases were considered to have
233 sufficient tissue for RNA extraction. The majority of cases (91%, 122 of 134) yielded
234 RNA of suitable quality and quantity for the NanoString assay. All cases were
235 successfully analysed using this assay, and the raw data generated passed the
236 relevant quality control parameters. The training set (n = 56) was comprised of 20
237 cases of OPMD that underwent malignant transformation (MT) and 36 cases that

238 were non-transforming (NT). The clinico-pathological features of the training set are
239 shown in Table 1. The test set (n = 66) was made up of 23 MT and 43 NT cases.
240 The clinico-pathological features of the test set are shown in Table 2. All the OPMD
241 in the study had oral epithelial dysplasia. Kaplan-Meier time to event analyses (time
242 to malignant transformation) for low- and high-grade epithelial dysplasia are shown
243 in Supplementary Figure 1 for both the training and the test sets. An accompanying
244 swimmer plot of the timing of individual events and censor dates are presented in
245 Supplementary Figure 2.
246 Following pre-processing and normalisation of the NanoString gene expression data
247 (Materials and Methods), univariable prognostic association of genes in the training
248 and test sets was assessed. Of the 33 genes, eight were prognostic in the training
249 set (Wald $p < 0.05$; Supplementary Table 2) and five were prognostic in the test set
250 (Wald $p < 0.05$; Supplementary Table 3). Three genes (*NOTCH1*, *CD274* and
251 *ITGB8*) were prognostic in both sets and also demonstrated consistency in the
252 direction of the estimated hazard ratio (Training set: *NOTCH1* HR = 0.26 & p =
253 0.009, *CD274* HR = 2.76 & p = 0.032, *ITGB8* HR = 3.04 & p = 0.023; Test set:
254 *NOTCH1* HR = 0.19 & p = 6.7×10^{-4}, *CD274* HR = 4.81 & p = 0.001, *ITGB8* HR =
255 5.55 & p = 0.002; Supplementary Figure 3). Lower *NOTCH1* transcription and higher
256 levels of *CD274* and *ITGB8* transcripts were associated with malignant progression.
257 Next, we used the training set to identify a prognostic gene signature associated with
258 malignant transformation. A multivariable prognostic model (Cox model with *L1*-
259 regularization; 4-fold cross validation) was created which constituted 11 genes. The
260 gene list together with the relevant weightage of each gene is shown in Table 3. This
261 prognostic model was used to predict patient risk scores in the test set. The
262 predicted risk scores were dichotomised into high- and low-risk groups (using
263 median risk score derived from the training set). The risk groups demonstrated two
264 clusters of patients in the test set when assessed against the mRNA abundance data
265 of the underlying genes in the multivariable prognostic model (Figure 1). These risk
266 groups were further tested for association with patient outcome using Cox
267 proportional hazards model adjusting for age at diagnosis, sex, site, type of OPMD
268 and binary OED grade. The prognostic gene signature remained an independent
269 predictor of malignant transformation when assessed in the test set, with high-risk
270 group showing worse prognosis (hazard ratio (HR) = 12.65, p = 0.0003; Figure 2a
271 and Table 4). In the multivariable setting, in addition to the gene signature-derived

272 risk scoring, binary OED grading was also statistically significant (p = 0.017).
273 Predicted risk groups were also tested for association with malignant transformation
274 using C-index which also confirmed strong concordance between the predicted risk
275 groups and survival times (Concordance index = 0.82, 0.75 – 0.88).
276 The predicted risk groups were verified for potential bias in the expression of the
277 housekeeping genes (*GAPDH, SDHA, TBP, TUBB*), which showed stable
278 expression levels across both groups except for a nominal difference in TUBB
279 expression in the test set (log2 fold change = 0.23, p = 0.01, Wilcoxon rank-sum test)
280 (Supplementary Figure 4).
281 Although our predictor was trained and tested using the Nanostring nCounter
282 platform, we tested it in an external cohort (GSE26549),[26] which was profiled using a
283 microarray platform (Supplementary Methods). Our classifier accurately predicted
284 the risk of oral cancer free survival in this independent cohort (HR = 2.38, p = 0.041)
285 despite the differences arising from the RNA quantifying platform (Figure 2b).
286 Furthermore, we used the gene signature to explore the association with normal and
287 malignant states in another microarray profiled dataset (GSE9844).[41] We observed
288 significantly elevated risk scores in tongue squamous cell carcinoma samples
289 compared to normal oral mucosa samples confirming oncogenic roles of the
290 signature genes exclusive to tumour samples ($p = 3.2 \times 10^{-5}$, Wilcoxon rank-sum
291 test, Figure 2c and Supplementary Methods).
292

293 **Discussion**

294 Currently, risk-stratification of OPMD patients in clinical practice is usually based on
295 a combination of clinical and histopathological features.[1, 23, 42] However, the
296 prognostic utility of these features has been found to be lacking and inconsistent.[1, 11,
297 15, 43] In this study, when considering clinico-pathological parameters, only OED
298 grading was found to be statistically significant in the training set. When the clinico-
299 pathological variables were fitted together using a Cox proportional hazards model,
300 only the binary OED grading of cases was found to be statistically significant. This
301 suggests that of all the clinico-pathological parameters, OED grading is the most
302 useful prognostic indicator for malignant transformation in OPMD and supports the
303 use of the binary grading system in clinical practice. This is consistent with the
304 findings of most studies that have indicated that OED grading is currently the 'gold-
305 standard' for prognosticating clinical outcome in OPMD cases.[1, 23] A confounding

306 factor in the accurate risk assessment of the patients in this study was the lack of
307 data on smoking habits. Smoking status is typically presented in broad categories
308 such as current smoker, ex-smoker, never smoker, however, there are very few
309 studies that provide detailed life-time exposure in pack-years, furthermore there is
310 evidence that patients tend to under-report their smoking habits leading to inaccurate
311 risk estimates.[44]
312 Archived FFPE tissues are an invaluable resource that can be successfully used for
313 molecular-based assays despite the degradation of nucleic acids that accompanies
314 fixation and embedding of samples in paraffin wax.[27, 28, 45-48] Our study provides
315 evidence of the clinical utility of the NanoString nCounter platform in providing robust
316 gene expression outputs using RNA from FFPE tissue.[27, 28, 30, 31, 49, 50] Though
317 relatively new, the NanoString nCounter assay has been shown by several studies to
318 be sensitive and reproducible, with sensitivity and accuracy levels that are better
319 than microarrays and comparable to real-time quantitative PCR (qPCR).[28, 30, 31] A
320 recent study by Veldman-Jones et al (2015) that evaluated the robustness of the
321 nCounter platform in analysing clinical samples showed that the platform has high
322 sensitivity of target detection and good reproducibility even with low RNA amounts,
323 making it suitable for developing clinical tests.[30] There are two main advantages of
324 NanoString technology compared to conventional gene expression analysis methods
325 such as qPCR and microarrays. In the nCounter platform, transcript levels are
326 measured from non-amplified total RNA, unlike other platforms, thus reducing
327 errors/biases that may be introduced through increased sample manipulation and
328 enzymatic reactions.[27, 28] Another advantage of NanoString is that it can be
329 multiplexed to measure up to 800 target genes in one sample, unlike qPCR-based
330 methods that are usually only able to measure the expression of a few genes at a
331 time.[27, 28, 30, 31] These features were key to developing Prosigna™, which is a
332 licenced prognostic test for breast cancer.
333 The gene signature developed in our study using the NanoString assay shows good
334 potential in prognosticating clinical outcome. Our findings are analogous to the
335 findings reported by Saintigny et al (2011) where the authors showed that gene
336 expression-based methods were superior to clinical and histological variables in
337 determining clinical outcome in OPMD patients.[26] In their study, Saintigny et al
338 (2011) compared microarray-derived gene expression-based models against a
339 model that contained only age, histology (dysplasia vs hyperplasia) and two

340 biomarkers (ΔNp63 and podoplanin).[26] The two models containing microarray data
341 showed much better performance compared to the model without any microarray
342 data. Their final model, which combined the microarray data with clinical and
343 pathological covariates, showed a slight improvement compared to the model with
344 only microarray data. However, only nine transcripts were similar between the two
345 microarray-based models, highlighting the rather unstable methodology employed in
346 constructing their prognostic model. Aside from that, other major differences
347 between their study and the current study are the type of tissue, the platform utilised
348 to obtain the gene expression data and the statistical methodology used to arrive at
349 the final gene expression profile.[26] Nevertheless, our gene classifier accurately
350 predicted the risk of oral cancer free survival in the Saintigny dataset.[26] We also
351 discovered that our gene signature was significantly different in matched normal oral
352 mucosa samples and tongue squamous cell carcinoma.[41] Together, these data
353 suggest that the gene expression-derived classifier reported in this study is
354 potentially generalisable and is likely to be underpinned by biologically relevant
355 changes in oral carcinogenesis. Several novel genes (*TLX1, CCNE1, ITGB8* and
356 *COL4A5*) with no known prior associations with oral carcinogenesis contributed to
357 the gene signature that was developed. The characteristics of all the genes in the
358 classifier are summarised in Table 3.[51-75]
359 One major issue with prognostic/predictive models is clinical validation. For example,
360 the molecularly driven prognostic model for malignant transformation of oral
361 leukoplakia developed by Saintigny et al (2011), though initially promising, has not
362 been translated into clinical practice.[26] To promote translation into clinical practice,
363 new prognostic/predictive models should be validated by an independent research
364 team using independent patient cohorts.[76] Lack of independence between the
365 training and test/validation cohorts can lead to an over-estimation of the prognostic
366 ability of such models. Another barrier for successful validation of a prognostic gene
367 signature is the presence of inter- and intra-tumour heterogeneity in OSCC, as well
368 as heterogeneity in OPMD.
369 Even though the current study has demonstrated the value of a molecularly driven
370 prognostic model over traditional risk-stratification methods for OPMD patients,
371 molecular-based methods are not without their drawbacks. A major limitation of the
372 current study is the sample size and the almost equal number of MT and NT cases
373 that is not truly representative of the population where MT is variable and ranges

between 0.13-36.4% depending on the cohort.[9] However, this study was designed to be a proof-of-principle study to explore the possibility of using FFPE-derived material for development of a gene signature prognostic of clinical outcome in OPMD patients. As such, we acknowledge that our study is only the first step in the development of a definitive gene expression-based prognostic model for OPMD. We also recognise that Nanostring is an expensive ‘research use only’ assay, nevertheless, it is conceivable that development of a clinical test would reduce costs by economy of scale. Prosigna™ a Nanostring-based breast cancer test, is proof that the technology can be translated into a cost-effective clinical test.

Although our study has successfully shown that the prognostic model developed is superior to conventional risk-stratification methods in a test set, the patients were obtained in a retrospective manner and the number of samples was small. Future studies require external validation in a sufficiently powered, prospective cohort study, recruiting consecutive patients with OPMD or as an observational component in a clinical trial. Ideally, such studies should be large enough to allow for data to be analysed by dysplasia grade, since this would provide valuable insight into the strengths and limitations of the molecular classifier against the current gold standard for risk assessment.

Conclusions

We have shown proof of principle that RNA extracted from FFPE-tissue, when analysed on the NanoString nCounter platform, can be used to model a gene expression signature that accurately predicts the risk of oral potentially malignant disorders undergoing malignant transformation. The molecular classifier was developed on a training set and validated on a test set, but still requires external validation in an appropriately powered cohort study before it can be used in clinical practice.

Additional information

Acknowledgements

We would like to acknowledge Ms. Anastasia Resteu from the Newcastle NanoString Unit, Newcastle University, for assisting with the NanoString assay.

408 **Authors' contributions**
409 HPS conceived and designed the study, performed experiments, analysed data,
410 carried out statistical analysis and wrote major parts of the manuscript. RK
411 participated in study design, provided expert advice, performed experiments,
412 interpreted results and edited the manuscript. PS participated in study design,
413 provided expert advice, performed dysplasia grading, contributed to retrieval of
414 tissue, interpreted results and edited the manuscript. PT and MN contributed to case
415 selection, retrieval of follow up data, quality assurance of data, interpreted results
416 and edited the manuscript. JA and SH contributed to study design, analysed data,
417 carried out statistical analysis, quality assurance of results, interpreted results and
418 wrote major parts of the manuscript. MR conceived and designed the study, provided
419 expert advice, performed dysplasia grading, contributed to retrieval of tissue,
420 analysed data, interpreted results and wrote major parts of the manuscript. All
421 authors read and approved the final manuscript.
422
423 **Ethics approval and consent to participate**
424 The study was performed with approval from a Health Research Authority (UK)
425 Research Ethics Committee (North East - Tyne and Wear South Research Ethics
426 Committee; NRES Committee North East – Sunderland 11/NE/0118) and complies
427 with UK legislation and guidelines. Patients were not recruited to the study and
428 therefore individual patient consent was not sought. Link-anonymised patient tissue
429 samples, surplus to diagnostic requirements, were analysed in accordance with the
430 terms of the ethical approval. The study was performed in accordance with the
431 Declaration of Helsinki.
432
433 **Consent for publication**
434 Not applicable.
435
436 **Data availability**
437 The datasets used and/or analysed during the current study are available at
438 https://doi.org/10.5281/zenodo.4643470.
439
440 **Competing interests**
441 The authors declare no conflicts of interest.

442
443 **Funding information**
444 Hans Prakash Sathasivam was supported by scholarship from Ministry of Health
445 Malaysia. Syed Haider is supported by Breast Cancer Now as part of Programme
446 Funding to The Breast Cancer Now Toby Robins Research Centre.
447
448 **References**

449 1. El-Naggar AK, Chan JKC, Grandis JR, Takata T, Slootweg PJ, editors. WHO
450 Classification of Head and Neck Tumours. Lyon, France: International Agency for
451 Research on Cancer; 2017.

452 2. Barnes L, Eveson JW, Reichart P, Sidransky D, editors. Pathology and
453 Genetics of Head and Neck Tumours. Lyon, France: World Health Organization ,
454 IARC; 2005.

455 3. Moore S, Johnson N, Pierce A, Wilson D. The epidemiology of lip cancer: a
456 review of global incidence and aetiology. Oral Dis. 1999;5(3):185-95.

457 4. Moore SR, Johnson NW, Pierce AM, Wilson DF. The epidemiology of mouth
458 cancer: a review of global incidence. Oral Dis. 2000;6(2):65-74.

459 5. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer.
460 Oral Oncol. 2009;45(4-5):309-16.

461 6. Goodson ML, Thomson PJ. Management of oral carcinoma: benefits of early
462 precancerous intervention. Br J Oral Maxillofac Surg. 2011;49(2):88-91.

463 7. van der Waal I. Potentially malignant disorders of the oral and oropharyngeal
464 mucosa; terminology, classification and present concepts of management. Oral
465 Oncol. 2009;45(4-5):317-23.

466 8. Warnakulasuriya S, Johnson NW, van der Waal I. Nomenclature and
467 classification of potentially malignant disorders of the oral mucosa. J Oral Pathol
468 Med. 2007;36(10):575-80.

469 9. Warnakulasuriya S, Ariyawardana A. Malignant transformation of oral
470 leukoplakia: a systematic review of observational studies. J Oral Pathol Med.
471 2016;45(3):155-66.

472 10. Shariff JA, Zavras AI. Malignant Transformation Rate in Patients Presenting
473 Oral Epithelial Dysplasia: Systematic Review and Meta-Analysis. Journal of Oral
474 Diseases. 2015;2015:1-10.

475 11. Mehanna HM, Rattay T, Smith J, McConkey CC. Treatment and follow-up of
476 oral dysplasia - a systematic review and meta-analysis. Head Neck.
477 2009;31(12):1600-9.

478 12. Thomson PJ, Goodson ML, Smith DR. Profiling cancer risk in oral potentially
479 malignant disorders-A patient cohort study. J Oral Pathol Med. 2017;46(10):888-95.

480 13. Napier SS, Speight PM. Natural history of potentially malignant oral lesions
481 and conditions: an overview of the literature. J Oral Pathol Med. 2008;37(1):1-10.

482 14. Field EA, McCarthy CE, Ho MW, Rajlawat BP, Holt D, Rogers SN, et al. The
483 management of oral epithelial dysplasia: The Liverpool algorithm. Oral Oncol.
484 2015;51(10):883-7.

485 15. Villa A, Woo SB. Leukoplakia—A Diagnostic and Management Algorithm.
486 Journal of Oral and Maxillofacial Surgery. 2017;75(4):723-34.

487 16. Awadallah M, Idle M, Patel K, Kademani D. Management update of potentially
488 premalignant oral epithelial lesions. Oral Surg Oral Med Oral Pathol Oral Radiol.
489 2018;125(6):628-36.

490 17. Reibel J. Prognosis of oral pre-malignant lesions: significance of clinical,
491 histopathological, and molecular biological characteristics. Crit Rev Oral Biol M.
492 2003;14(1):47-62.

493 18. Pitiyage G, Tilakaratne WM, Tavassoli M, Warnakulasuriya S. Molecular
494 markers in oral epithelial dysplasia: review. J Oral Pathol Med. 2009;38(10):737-52.

495 19. Nankivell P, Mehanna H. Oral dysplasia: biomarkers, treatment, and follow-
496 up. Curr Oncol Rep. 2011;13(2):145-52.

497 20. Smith J, Rattay T, McConkey C, Helliwell T, Mehanna H. Biomarkers in
498 dysplasia of the oral cavity: a systematic review. Oral Oncol. 2009;45(8):647-53.

499 21. Lingen MW, Pinto A, Mendes RA, Franchini R, Czerninski R, Tilakaratne WM,
500 et al. Genetics/epigenetics of oral premalignancy: current status and future research.
501 Oral Dis. 2011;17 Suppl 1:7-22.

502 22. Nikitakis NG, Pentenero M, Georgaki M, Poh CF, Peterson DE, Edwards P, et
503 al. Molecular markers associated with development and progression of potentially
504 premalignant oral epithelial lesions: Current knowledge and future implications. Oral
505 Surg Oral Med Oral Pathol Oral Radiol. 2018;125(6):650-69.

506 23. Speight PM, Khurram SA, Kujan O. Oral potentially malignant disorders: risk
507 of progression to malignancy. Oral Surg Oral Med Oral Pathol Oral Radiol.
508 2018;125(6):612-27.

509 24. De Cecco L, Nicolau M, Giannoccaro M, Daidone MG, Bossi P, Locati L, et al.
510 Head and neck cancer subtypes with biological and clinical relevance: Meta-analysis
511 of gene-expression data. Oncotarget. 2015;6(11):9627-42.

512 25. Tonella L, Giannoccaro M, Alfieri S, Canevari S, De Cecco L. Gene
513 Expression Signatures for Head and Neck Cancer Patient Stratification: Are Results
514 Ready for Clinical Application? Curr Treat Options Oncol. 2017;18(5):32.

515 26. Saintigny P, Zhang L, Fan YH, El-Naggar AK, Papadimitrakopoulou VA, Feng
516 L, et al. Gene expression profiling predicts the development of oral cancer. Cancer
517 Prev Res (Phila). 2011;4(2):218-29.

518 27. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al.
519 Direct multiplexed measurement of gene expression with color-coded probe pairs.
520 Nat Biotechnol. 2008;26(3):317-25.

521 28. Reis PP, Waldron L, Goswami RS, Xu W, Xuan Y, Perez-Ordonez B, et al.
522 mRNA transcript quantification in archival samples using multiplexed, color-coded
523 probes. Bmc Biotechnol. 2011;11:46.

524 29. Balko JM, Cook RS, Vaught DB, Kuba MG, Miller TW, Bhola NE, et al.
525 Profiling of residual breast cancers after neoadjuvant chemotherapy identifies
526 DUSP4 deficiency as a mechanism of drug resistance. Nat Med. 2012;18(7):1052-9.

527 30. Veldman-Jones MH, Brant R, Rooney C, Geh C, Emery H, Harbron CG, et al.
528 Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter
529 Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples.
530 Cancer Res. 2015;75(13):2587-93.

531 31. Veldman-Jones MH, Lai Z, Wappett M, Harbron CG, Barrett JC, Harrington
532 EA, et al. Reproducible, Quantitative, and Flexible Molecular Subtyping of Clinical
533 DLBCL Samples Using the NanoString nCounter System. Clin Cancer Res.
534 2015;21(10):2367-78.

535 32. Speight PM, Abram TJ, Floriano PN, James R, Vick J, Thornhill MH, et al.
536 Interobserver agreement in dysplasia grading: toward an enhanced gold standard for
537 clinical pathology trials. Oral Surg Oral Med Oral Pathol Oral Radiol.
538 2015;120(4):474-82 e2.

539 33. Kujan O, Oliver RJ, Khattab A, Roberts SA, Thakker N, Sloan P. Evaluation of
540 a new binary system of grading oral epithelial dysplasia for prediction of malignant
541 transformation. Oral Oncol. 2006;42(10):987-93.

542 34. NanoString. nCounter XT Assay User Manual. Seattle, Washington USA:
543 NanoString Technologies Inc.; 2016.

544 35. Sathasivam HP, Casement J, Bates T, Sloan P, Thomson P, Robinson M, et
545 al. Gene expression changes associated with malignant transformation of oral
546 potentially malignant disorders. Journal of Oral Pathology & Medicine. 2020;n/a(n/a).

547 36. Lohavanichbutr P, Méndez E, Holsinger FC, Rue TC, Zhang Y, Houck J, et al.
548 A 13-Gene Signature Prognostic of HPV-Negative OSCC: Discovery and External
549 Validation. Clinical Cancer Research. 2013;19(5):1197.

550 37. Lallemant B, Evrard A, Combescure C, Chapuis H, Chambon G, Raynal C, et
551 al. Reference gene selection for head and neck squamous cell carcinoma gene
552 expression studies. BMC Mol Biol. 2009;10:78-.

553 38. Taïhi I, Nassif A, Berbar T, Isaac J, Berdal A, Gogly B, et al. Validation of
554 Housekeeping Genes to Study Human Gingival Stem Cells and Their <i>In Vitro</i>
555 Osteogenic Differentiation Using Real-Time RT-qPCR. Stem Cells International.
556 2016;2016:6261490.

557 39. Rentoft M, Hultin S, Coates PJ, Laurell G, Nylander K. Tubulin α-6 chain is a
558 stably expressed reference gene in archival samples of normal oral tissue and oral
559 squamous cell carcinoma. Exp Ther Med. 2010;1(3):419-23.

560 40. Haider S, Yao CQ, Sabine VS, Grzadkowski M, Stimper V, Starmans MHW,
561 et al. Pathway-based subnetworks enable cross-disease biomarker discovery. Nat
562 Commun. 2018;9(1):4746.

563 41. Ye H, Yu T, Temam S, Ziober BL, Wang J, Schwartz JL, et al. Transcriptomic
564 dissection of tongue squamous cell carcinoma. BMC genomics. 2008;9:69.

565 42. van der Waal I. Potentially malignant disorders of the oral and oropharyngeal
566 mucosa; present concepts of management. Oral Oncol. 2010;46(6):423-5.

567 43. Dost F, Le Cao K, Ford PJ, Ades C, Farah CS. Malignant transformation of
568 oral epithelial dysplasia: a real-world evaluation of histopathologic grading. Oral Surg
569 Oral Med Oral Pathol Oral Radiol. 2014;117(3):343-52.

570 44. Hald J, Overgaard J, Grau C. Evaluation of objective measures of smoking
571 status--a prospective clinical study in a group of head and neck cancer patients
572 treated with radiotherapy. Acta Oncol. 2003;42(2):154-9.

573 45. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA
574 quality from FFPE samples. PLoS One. 2007;2(12):e1261.

575 46. Eikrem O, Beisland C, Hjelle K, Flatberg A, Scherer A, Landolt L, et al.
576 Transcriptome Sequencing (RNAseq) Enables Utilization of Formalin-Fixed, Paraffin-
577 Embedded Biopsies with Clear Cell Renal Cell Carcinoma for Exploration of Disease
578 Biology and Biomarker Development. PLoS One. 2016;11(2):e0149743.

579 47. Mittempergher L, de Ronde JJ, Nieuwland M, Kerkhoven RM, Simon I,
580 Rutgers EJ, et al. Gene expression profiles from formalin fixed paraffin embedded
581 breast cancer tissue are largely comparable to fresh frozen matched tissue. PLoS
582 One. 2011;6(2):e17163.

583 48. Wimmer I, Troscher AR, Brunner F, Rubino SJ, Bien CG, Weiner HL, et al.
584 Systematic evaluation of RNA quality, microarray data reliability and pathway
585 analysis in fresh, fresh frozen and formalin-fixed paraffin-embedded tissue samples.
586 Sci Rep. 2018;8(1):6351.

587 49. Scott DW, Chan FC, Hong F, Rogic S, Tan KL, Meissner B, et al. Gene
588 expression-based model using formalin-fixed paraffin-embedded biopsies predicts
589 overall survival in advanced-stage classical Hodgkin lymphoma. J Clin Oncol.
590 2013;31(6):692-700.

591 50. Saba NF, Wilson M, Doho G, DaSilva J, Benjamin Isett R, Newman S, et al.
592 Mutation and Transcriptional Profiling of Formalin-Fixed Paraffin Embedded
593 Specimens as Companion Methods to Immunohistochemistry for Determining
594 Therapeutic Targets in Oropharyngeal Squamous Cell Carcinoma (OPSCC): A Pilot
595 of Proof of Principle. Head Neck Pathol. 2015;9(2):223-35.

596 51. Scaltriti M, Eichhorn PJ, Cortes J, Prudkin L, Aura C, Jimenez J, et al. Cyclin
597 E amplification/overexpression is a mechanism of trastuzumab resistance in HER2+
598 breast cancer patients. Proc Natl Acad Sci U S A. 2011;108(9):3761-6.

599 52. Cancer Genome Atlas N. Comprehensive genomic characterization of head
600 and neck squamous cell carcinomas. Nature. 2015;517(7536):576-82.

601 53. Hatano M, Roberts CW, Minden M, Crist WM, Korsmeyer SJ. Deregulation of
602 a homeobox gene, HOX11, by the t(10;14) in T cell leukemia. Science.
603 1991;253(5015):79-82.

604 54. Yap LF, Lee D, Khairuddin A, Pairan MF, Puspita B, Siar CH, et al. The
605 opposing roles of NOTCH signalling in head and neck cancer: a mini review. Oral
606 Dis. 2015;21(7):850-7.

607 55. Ranganathan P, Weaver KL, Capobianco AJ. Notch signalling in solid
608 tumours: a little bit of everything but not all the time. Nat Rev Cancer.
609 2011;11(5):338-51.

610 56. Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, et al.
611 Exome sequencing of head and neck squamous cell carcinoma reveals inactivating
612 mutations in NOTCH1. Science. 2011;333(6046):1154-7.

613 57. Alevizos I, Mahadevappa M, Zhang X, Ohyama H, Kohno Y, Posner M, et al.
614 Oral cancer in vivo gene expression profiling assisted by laser capture
615 microdissection and microarray analysis. Oncogene. 2001;20(43):6196-204.

616 58. Macabeo-Ong M, Shiboski CH, Silverman S, Ginzinger DG, Dekker N, Wong
617 DT, et al. Quantitative analysis of cathepsin L mRNA and protein expression during
618 oral cancer progression. Oral Oncol. 2003;39(7):638-47.

619 59. Kang C-J, Chen Y-J, Liao C-T, Wang H-M, Chang JT, Lin C-Y, et al.
620 Transcriptome profiling and network pathway analysis of genes associated with
621 invasive phenotype in oral cancer. Cancer Letters. 2009;284(2):131-40.

622 60. Mallery SR, Zwick JC, Pei P, Tong M, Larsen PE, Shumway BS, et al. Topical
623 Application of a Bioadhesive Black Raspberry Gel Modulates Gene Expression and
624 Reduces Cyclooxygenase 2 Protein in Human Premalignant Oral Lesions. Cancer
625 Research. 2008;68(12):4945.

626 61. Kuriakose MA, Chen WT, He ZM, Sikora AG, Zhang P, Zhang ZY, et al.
627 Selection and validation of differentially expressed genes in head and neck cancer.
628 Cellular and Molecular Life Sciences CMLS. 2004;61(11):1372-83.

629 62. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head
630 and neck cancer. Nat Rev Cancer. 2011;11(1):9-22.

631 63. Brennan JA, Boyle JO, Koch WM, Goodman SN, Hruban RH, Eby YJ, et al.
632 Association between cigarette smoking and mutation of the p53 gene in squamous-
633 cell carcinoma of the head and neck. N Engl J Med. 1995;332(11):712-7.

634 64. Apostolou P, Papasotiriou I. Current perspectives on CHEK2 mutations in
635 breast cancer. Breast Cancer (Dove Med Press). 2017;9:331-5.

636 65. Cybulski C, Górski B, Huzarski T, Masojć B, Mierzejewski M, Debniak T, et al.
637 CHEK2 is a multiorgan cancer susceptibility gene. Am J Hum Genet.
638 2004;75(6):1131-5.

639 66. Cybulski C, Huzarski T, Górski B, Masojć B, Mierzejewski M, Debniak T, et al.
640 A novel founder CHEK2 mutation is associated with increased prostate cancer risk.
641 Cancer Res. 2004;64(8):2677-9.

642 67. Chen Y-L, Tsai W-H, Ko Y-C, Lai T-Y, Cheng A-J, Shiah S-G, et al. Discoidin
643 Domain Receptor-1 (DDR1) is Involved in Angiolymphatic Invasion in Oral Cancer.
644 Cancers (Basel). 2020;12(4):841.

645 68. Wang WW, Wang YB, Wang DQ, Lin Z, Sun RJ. Integrin beta-8 (ITGB8)
646 silencing reverses gefitinib resistance of human hepatic cancer HepG2/G cell line. Int
647 J Clin Exp Med. 2015;8(2):3063-71.

648 69. Cui Y, Wu F, Tian D, Wang T, Lu T, Huang X, et al. miR-199a-3p enhances
649 cisplatin sensitivity of ovarian cancer cells by targeting ITGB8. Oncol Rep.
650 2018;39(4):1649-57.

651 70. Mertens-Walker I, Fernandini BC, Maharaj MS, Rockstroh A, Nelson CC,
652 Herington AC, et al. The tumour-promoting receptor tyrosine kinase, EphB4,
653 regulates expression of integrin-beta8 in prostate cancer cells. BMC Cancer.
654 2015;15:164.

655 71. Lenouvel D, González-Moles MÁ, Talbaoui A, Ramos-García P, González-
656 Ruiz L, Ruiz-Ávila I, et al. An update of knowledge on PD-L1 in head and neck
657 cancers: Physiologic, prognostic and therapeutic perspectives. Oral Diseases.
658 2020;26(3):511-26.

659 72. Dave K, Ali A, Magalhaes M. Increased expression of PD-1 and PD-L1 in oral
660 lesions progressing to oral squamous cell carcinoma: a pilot study. Scientific
661 Reports. 2020;10(1):9705.

662 73. Saintigny P, El-Naggar AK, Papadimitrakopoulou V, Ren H, Fan YH, Feng L,
663 et al. DeltaNp63 overexpression, alone and in combination with other biomarkers,
664 predicts the development of oral cancer in patients with leukoplakia. Clin Cancer
665 Res. 2009;15(19):6284-91.

666 74. Matsubara R, Kawano S, Kiyosue T, Goto Y, Hirano M, Jinno T, et al.
667 Increased DeltaNp63 expression is predictive of malignant transformation in oral
668 epithelial dysplasia and poor prognosis in oral squamous cell carcinoma.
669 International journal of oncology. 2011;39(6):1391-9.

670 75. Chen YK, Hsue SS, Lin LM. Expression of p63 protein and mRNA in oral
671 epithelial dysplasia. J Oral Pathol Med. 2005;34(4):232-9.

672 76. Chibon F. Cancer gene expression signatures - the rise and fall? Eur J
673 Cancer. 2013;49(8):2000-9.

674
675

676 **Figure legends**
677 **Figure 1.** Bar plot showing the predicted risk scores of the samples in the test set.
678 The clinico-pathological covariates age at diagnosis, sex, OED grade, type of OPMD
679 and site of the index lesion are shown in rows below the bar plot. A heat map shows
680 the mRNA abundance (z-score) of the genes from the prognostic gene signature for
681 the test set samples.
682
683 **Figure 2.** Kaplan-Meier time to event analysis using Cox proportional hazards model
684 comparing malignant transformation in the test set samples divided into low- or high-
685 risk groups. Predicted risk scores were categorised into low- and high-risk groups
686 using a threshold estimated as the median risk score of the training set (a). The gene
687 expression-derived classifier was informative in an independent cohort (GSE26549
688 dataset)[26] (b) and was biologically relevant as the predicted risk scores were
689 significantly higher in tongue squamous cell carcinoma samples compared with
690 normal oral mucosa samples (GSE9844 dataset;[41] Wilcoxon rank-sum test) (c).
691
692 **Table legends**
693 **Table 1.** Clinico-pathological features of training set (n = 56).
694
695 **Table 2.** Clinico-pathological features of test set (n = 66).
696
697 **Table 3.** Prognostic gene signature for patients with OPMD along with the estimated
698 beta coefficients (weightage).
699
700 **Table 4.** Multivariable Cox proportional hazards model (test set).