Institute of Head and Neck Studies and Education

UNIVERSITY OF BIRMINGHAM

| | |
|---|---|
| Full protocol title: | **Validating the PredicTR treatment response classifier for oropharyngeal cancer** |
| Acronym: | **PredicTR 2** |
| Version: | **1.0** |
| Date: | **08 October 2018** |
| Sponsor: | University of Birmingham |
| Sponsor Reference: | RG_18-176 |
| Funding Body: | MRC BMC:DPFS |

## Contents

## Contact names and numbers

**Chief Investigator**

Prof Hisham Mehanna
Professor of Head and Neck Surgery
Institute of Head and Neck Studies and Education (InHANSE)
Institute of Cancer and Genomic Sciences
Robert Aitken Building, 2nd Floor
University of Birmingham
Birmingham. B15 2TT
E-mail: h.mehanna@bham.ac.uk
Tel: 0121 414 6547

Signature: 
Date: 08/10/2018

**Project Manager**

Dr Rachel Spruce
Translational Team Leader
Institute of Head and Neck Studies and Education (InHANSE)
Institute of Cancer and Genomic Sciences
Robert Aitken Building, 2nd Floor
University of Birmingham
Birmingham. B15 2TT
E-mail: r.spruce@bham.ac.uk
Tel: 0121 414 9275

**Sponsor contact**

Research Support Group
University of Birmingham
The Dome, Aston Webb, C Block
Birmingham
United Kingdom
B15 2TT
Email: researchgovernance@contacts.bham.ac.uk

Note: Signature on the IRAS form by the Sponsor constitutes sponsorship agreement, and therefore is not required on the protocol.

**Co-Investigators and Project Partners**

Mr Stuart Winter
Blenheim Head and Neck Unit
Churchill Hospital
Old Road, Headington
Oxford, OX3 7LE
Email: stuart.winter@ouh.nhs.uk
Tel: 01865 231 061

Dr Max Robinson
Senior Lecturer in Oral Pathology
School of Dental Sciences
Newcastle University
Framlington Place
Newcastle-upon-Tyne, NE2 4BW
Email: max.robinson@ncl.ac.uk
Tel: 0191 208 8391

Professor Andy Ness
NIHR Bristol Biomedical Research Centre Nutrition Theme
Level 3, University Hospitals Bristol Education Centre
Upper Maudlin Street
Bristol, BS2 8AE
Email: andy.ness@bristol.ac.uk
Tel: 01173421751

Professor Steven Thomas
Division of Oral and Maxillofacial Surgery
Bristol Dental Hospital
Lower Maudlin Street
Bristol, BS1 2LY
Email: Steve.Thomas@bristol.ac.uk

Dr Miranda Pring
Bristol Dental Hospital
Lower Maudlin Street
Bristol, BS1 2LY
Email: Miranda.Pring@bristol.ac.uk

Anthony Kong
Institute of Head and Neck Studies and Education (InHANSE)
University of Birmingham
Robert Aitken Building, 2nd Floor
Birmingham, B15 2TT
Email: A.H.Kong@bham.ac.uk
Tel: 0121 414 9294

Mr Paul Nankivell
Senior Clinical Lecturer
Institute of Head and Neck Studies and Education (InHANSE)
University of Birmingham
Robert Aiken Building, 2nd Floor
Birmingham, B15 2TT
Email: P.C.Nankivell@bham.ac.uk
Tel: 0121 414 6547

Professor Nasir Rajpoot
Department of Computer Science
University of Warwick,
Coventry, CV4 7AL
Email: N.M.Rajpoot@warwick.ac.uk
Tel: 07698 0999

Dr Ketan Shah
Cellular Pathology
The John Radcliffe Hospital
Oxford, OX3 9DU
Email: ketan.shah@nhs.net
Tel: 01865 222893

Dr Jacqueline James
Northern Ireland Molecular Pathology Laboratory
Centre for Cancer Research and Cell Biology
Queens University Belfast
97 Lisburn Road
Belfast, BT9 7BL
Email: j.james@qub.ac.uk

Professor Gareth Thomas
University of Southampton
University Road
Southampton, SO17 1BJ
United Kingdom
Email: g.thomas@soton.ac.uk

Dr Selavam Thavaraj
Head & Neck Pathology
4th Floor Tower Wing, Guy's Hospital
Great Maze, Pond
London SE1 9RT UK
Email: selvam.thavaraj@kcl.ac.uk
Tel: 020 7188 4367

**Statisticians**

Professor Christopher Holmes
Department of Statistics
University of Oxford
24-29 St Giles'
Oxford, OX1 3LB
E-mail: cholmes@stats.ox.ac.uk
Tel: 01865 285873

Dr Jennifer Rogers
Department of Statistics
University of Oxford
24-29 St Giles'
Oxford, OX1 3LB
E-mail: jennifer.rogers@stats.ox.ac.uk
Tel: 07729 984952

Professor Jon Deeks
Institute of Applied Health Research
College of Medical and Dental Sciences
University of Birmingham
Birmingham, B15 2TT
E-mail: J.Deeks@bham.ac.uk
Tel: 0121 4145 328

## Executive Summary

We have undertaken preclinical development and internal validation on a retrospective cohort of our PredicTR biomarker classifier, comprising p16, HPV ISH, Survivin and Tumour Infiltrating Lymphocyte score (TILs). This classifier strongly predicts which oropharyngeal cancer (OPC) patients will benefit from adding surgery to current standard of care chemo-radiotherapy (20% improvement in 3-year overall survival (OS) for high risk patients), so could be used to decide treatment.

To progress validation along the NCI-Translational Research Working Group Developmental Pathway, and prepare for a definitive clinical trial, we aim to:

1. Optimise and implement the CE-marked Survivin reagent into our classifier. We will demonstrate equivalence with the Research Use Only reagent used during classifier development, by staining 200 samples from the previously tested cohort, in a bridging study.

2. Assess reproducibility and prognostic performance of an automated algorithm for scoring TILs in OPC. We developed an AI-based algorithm for scoring TILs on digitised images of H&E stained slides in oral cancer. We will optimise then compare its ability to predict OS versus scoring by pathologists on 96 OPC samples.

3. Test assay intra- and inter-laboratory reproducibility for the 4 biomarkers, using CE-marked reagents, on whole tissue sections in routine NHS clinical laboratories. In ring experiments, at least 25 randomised reference cases will be stained blind, in duplicate, in up to 6 separate NHS labs, and scored by trained pathologists blinded to outcome. If poorly reproducible, a qualitative study and further experiments will be done to address causes.

4. Validate the biomarker classifier in a prospective, independent external cohort under routine clinical conditions. Approximately 504 OPC samples, already collected prospectively within the Head and Neck 5000 cohort, will be stained and scored blind at up to 6 NHS laboratories. Classifier performance and prediction accuracy will be evaluated.

## Background

Chemo-radiotherapy (CRT, chemotherapy + radiotherapy given together) is the standard treatment for oropharyngeal cancer (OPC). With the recent advent of transoral surgery, there is uncertainty regarding the role of adding surgery or substituting it (instead of chemotherapy) in the 70% patients who are respectable to improve outcomes.

Adding surgery may improve outcomes in those patients who are at increased risk of locoregional recurrence. Surgery also eliminates the need for chemotherapy in ~ 50% patients, i.e. patients are treated with surgery and radiotherapy only, reducing risk of life-threatening infections, dry mouth and difficulty swallowing. However, the remaining 50% receive triple therapy (surgery + radiotherapy + chemotherapy), which comes with potentially increased complications (bleeding, infection), poorer long-term functional outcomes (dysphagia, aspiration, permanent tracheostomy) and increased cost (+~£5000).

Recent research has developed prognostic classifiers and nomograms that predict better prognosis regardless of treatment type, but cannot guide treatment selection. Most are based on HPV status, clinical stage (Tumour and Nodal stage), and smoking history (1-3). These novel approaches are currently being used to stratify and recruit patients into clinical trials, but they do not guide selection of specific treatments. Therefore, presently treatment decisions are based only on clinician preference and patient choice.

The concept of predictive classifiers (that can actually guide treatment selection) is well established (e.g. OncotypeDx in breast cancer). To date, there are no validated predictive classifiers for OPC.

## Development

Through systematic review and meta-analysis of literature, we identified 10 biomarkers that showed prognostic ability for OPC specifically (4). Following the NCI's Translational Working Group Biospecimen-based Developmental Pathway (NCI-TWGBDP), we optimised assays for these 10 biomarkers. The assays achieved good reproducibility and demonstrated high correlation >0.75 between blinded pathologists and between sections and tissue microarrays (TMAs).

In a single research lab, we then stained and scored the 10 biomarkers on TMAs of a development multi-centre cohort (n=600 retrospectively collected cases). We ensured adequate blinding of assessments and analyses, and adjusted for important prognostic factors.

Using Cox proportional hazards models, we developed a prognostic biomarker model that also predicts which OPC patients benefit from surgery. During development, the biomarker only model outperformed models with clinical factors only (with either TNM 7 or 8), and those with clinical factors and biomarkers combined.

**PredicTR Classifier Composition**

The biomarker-only model comprises p16 immunohistochemistry (IHC), HPV-DNA in situhybridisation (HPV-ISH), Survivin IHC and Tumour Infiltrating Lymphocytes score (TILs), spanning HPV status, immune status and cell cycle biology. All biomarkers are commercially available.

p16 and HPV-ISH are markers of HPV infection, and are well-established prognostic biomarkers in OPC (5). They have been very well-validated and used extensively in OPC, with well-defined cut-offs (6,7). p16 is scored positive if moderate to strong staining in 70% or more of tumour cells is seen, and HPV-ISH is scored either positive or negative (6).

Tumour Infiltrating Lymphocyte score (TILs) is widely accepted as a surrogate indicator for immune response to cancer. TILs are scored on an H&E slide under 2.5x magnification, as high (diffuse, T-Lymphocytes present in >80% tumour/stroma), low (weak/absent staining, <20% tumour/stroma) and moderate (patchy, present in 20-80% tumour/stroma) (8). TILs has been found to be an independent prognostic marker of overall and progression free survival in OPC (9): Compared to TILs low, patients scored as TILs moderate or high demonstrated a hazards ratio of 0.59 (0.37-0.95) and 0.21(0.11-0.41, p<0.001) respectively. Significantly more HPV-positive patients were TIL high. Immunohistochemistry staining of T-lymphocytes using CD3, CD8, CD4, FOXP3, and ratios of the above did not improve prognostication over TILs.

Survivin is a member of the inhibitor of apoptosis (IAP) gene family, which encode regulatory proteins that prevent apoptosis. It is generally deregulated in cancer (10) and is known to be over-expressed in OPC (4). A recent meta-analysis of Survivin expression in oral cancer demonstrated that increased expression was associated with poor prognosis (Hazard ratio of death 1.62) (11). Survivin immunohistochemistry is scored by H-score.

**Validation and Performance**

We then validated our classifier on TMAs from an independent retrospective multicenter cohort (n=385), replicating predictive performance with very good discrimination, Harrell's C-index 0.84 (95% CI 0.79-0.88), and excellent calibration.

Our biomarker algorithm classifies patients into low and high-risk. Low-risk patients do not benefit from additional surgery (HR=0.85, 95% CI 0.31-2.35, p=0.759) and can be treated with CRT alone. High-risk patients demonstrate ~20% absolute improvement in overall survival if surgery is added (3yr OS 63% versus 42.5% CRT alone; HR= 0.51, 95%CI 0.3-0.85, p=0.01). This benefit would constitute a greater improvement in survival than that caused by any advance in HNC treatment in last 2 decades.

The biomarker panel is relatively easy to read by pathologists, demonstrating high correlation (>0.75) between pathologists following a short 3-hour calibration process. It is relatively inexpensive (actual cost £100, retail £500-750). It is also possible to stain

and score quickly in a timely manner (3 working days), fitting well within the current treatment pathway.

**Cost-Benefit**

In most countries where CRT is the standard of care, adding surgery can result in triple therapy for 50% of patients with the potential for increased complications during treatment period, additional hospital stay, long term functional deficits (including poor swallow, aspiration, and permanent tracheostomy) and reduced quality of life, as well as significant additional cost (+£5000 per case).

Conversely, including surgery in the initial treatment of those patients who are at high risk of loco-regional recurrence appears to result in ~20% improvement in absolute overall survival. Curing (and preventing recurrence) in an additional 20% of the high-risk patients would be cost effective, saving substantial palliative care costs and the emotional, social and economic costs of early death of a member of society.

Therefore, our test, which would cost £500-750 (retail) per patient, could result in the 50% of patients who are high-risk having additional surgery (~£5,000 per case), but would result in cure of an additional 20% of these patients. This would save ~£30,000 palliative treatment costs per patient. In the UK, on the basis of 2000 OPC patients a year, implementation of our classifier is calculated to save a net £4.25 million per year.

Alternatively, if adding surgery becomes standard of care as is currently happening in some countries e.g. Germany and USA, then our test (costing £500-750) would save the 50% of patients who are low risk from having additional surgery (£5000 per case), resulting in reduced costs (projected in UK to be £3.5 million per year), as well as reduced toxicity and long-term functional deficit and improved quality of life for those patients.

**Aim**

We now wish to complete our classifier's validation process, before undertaking a definitive clinical trial in preparation for clinical adoption. This would be the first validated biomarker classifier guiding treatment selection for OPC, and would present a significant advance in the personalised treatment of these patients.

**Experimental Design and Assays**

**Design**

To progress validation along the NCI-Translational Research Working Group Developmental Pathway, and prepare for a definitive clinical trial, we aim to:

1. Optimise and implement the CE-marked Survivin reagent into our classifier, in a bridging study.
2. Assess reproducibility and prognostic performance of an automated algorithm for scoring TILs in OPC.

3. Test assay intra- and inter-laboratory reproducibility for biomarkers, using CE-marked reagents, on whole tissue sections in routine NHS clinical laboratories.
4. External validation on a prospective cohort (see below for sample collections).

**Samples and Blinding**

Patient characteristics, outcomes and samples have already been prospectively collected. Biomarker samples will be stained and scored by laboratories and pathologists blinded to outcomes.

Samples have already been collected as part of the following studies:

1. Head and Neck 5000
2. PredicTR (original study)
3. PET NECK
4. AcceleraTED2
5. De-ESCALaTE Collect
6. CompARE Collect

**Assays**

Haematoxylin and eosin (H&E) stained sections from formalin-fixed, paraffin embedded tissue blocks are reviewed by a pathologist to confirm the diagnosis of squamous cell carcinoma, and to assign Tumour Infiltrating lymphocyte (TILs) score. Freshly cut 3µm tissue sections are used to perform immunohistochemistry (IHC; p16, Survivin) and high-risk HPV DNA in situ hybridization (HPV ISH). p16 (CINtec® p16 Histology, Roche Diagnostics) and HPV ISH (Ventana INFORMVIII Family 16 probe) are CE-marked kits, used in accordance with the manufacturer's instructions. For Survivin, the RUO antibody (clone EP2880Y, Abcam; used at 1:750 dilution following heat-induced epitope retrieval, pH6) will be compared with the CE-marked reagent (clone EP119, BioSB used as per manufacturer's protocol). The study will be carried out in accordance with Good Clinical Laboratory Practice (GCLP).

**Calibration and Scoring**

All pathologists undergo certification by attending a calibration meeting and scoring a test set before starting. TILs are manually scored on scanning magnification (x2.5 objective) and assigned one of the following categories: high TILs (diffuse; present in >80% of tumour/stroma), moderate TILs (patchy; present in 20-80% of tumour/stroma) or low TILs (weak/absent; present in <20% of tumour/stroma) (9). IHC is scored by assigning an intensity score (0, no staining; 1, weak; 2, moderate; 3, strong) and the percentage (0-100% in 5% increments) of malignant cells staining at each intensity. These parameters are combined to produce an H- score (product of intensity and percentage) from 0-300 (12). H-score for Survivin (continuous variable) is scaled and transformed to Z-score (13). p16 is dichotomised into positive and negative categories according to a previously described, clinically validated cut off (strong and diffuse nuclear and cytoplasmic staining present in ≥70% of the tumour; H-score equivalent ≥2 intensity x ≥70% = H score ≥140) (1,6). High risk HPV ISH is scored using a binary classification (positive vs. negative) (6).

## Methods, statistical analysis plan and formal sample size calculations

### Survivin bridging study

The Survivin antibody that we used in the previous study was not CE-marked. Since then, a CE-marked reagent is now available commercially. We will optimise this, and undertake a bridging study to ensure equivalent results by the CE marked version against the predicate test (Research Use Only assay).

Methods: The bridging study will involve staining approximately 200 cases on tissue microarrays using the new CE Marked Antibody in one laboratory, and will be scored by at least 2 pathologists independently. The results will then be compared with the results of the tissue microarrays previously stained by the RUO antibody.

The Altman and Bland 'limits of agreement' technique will be used to describe by how much the H-scores for the two different biomarkers differ. This approach compares measures of the same continuous quantity, and assesses systematic bias, trends in systematic bias with biomarker value, and the magnitude of differences between measures. These allow 95% "limits of agreement" to be calculated which summarise the magnitude of the differences.

We noted high levels of disagreement for the RUO assay in a previous interlaboratory evaluation study, where differences in the Survivin H-score between laboratories had standard deviations between 16 and 31. Due to this low reproducibility, we cannot expect close agreement between the RUO assay and the new assay. We will therefore judge a satisfactory replacement assay to have similar limits of agreement as the reproducibility study.

Sample size calculations for limits of agreement studies are based on the precision of the limits of agreement. Based on the standard error formula proposed by Bland and Altman, the precision of the estimate can be expressed in proportion to the standard deviation of the between test differences. Our proposed study of approximately 200 participants will provide adequate data such that the standard error of upper and lower 95% limits of agreement would be 12.5% of the standard deviation of the differences - so the standard error would be 3H units should the SD be 25H units. This precision is adequate to be able to judge comparability of the old Research use only and the new CE-marked survivin reagents.

### Reproducibility of biomarkers

Method: The reproducibility study will involve testing of each sample (to estimate intra laboratory variability SDa) twice (at two different times) in each of up to six laboratories (to estimate between laboratory variability SDi) from approximately 25 patients (to estimate between patient variability SDg); a total of approximately 300 sections for each biomarker.

Estimates will be obtained by fitting a random effects model (if necessary after log transformation of the data should it show positive skew). The consequences of

measurement (score) variability on classifier performance will be assessed via simulation using the previously available study data obtained during classifier development. The random effects model will allow us to model the intra- to inter lab variability and perform an ANOVA analysis of the contribution of each factor to the overall stability and variability of the measurement. It will also allow us to optimise the resulting combined score, down-weighting any factors that happen to exhibit poor reproducibility.

This sample size and study design has been created to ensure that estimates of the standard errors of estimates of SDa, SDi and SDg are made with adequate precision to draw inferences from coefficients of variation, the index of individuality and reference change values. The design ensures that SDa and SDi are estimated with the greatest precision as they describe the repeatability and reproducibility of measurements. For example, estimates of SDa, SDi and SDg for survivin from our previous reproducibility study were in the order of 10, 20 and 30 H; via simulation using 1000 replications these values will be estimated with 95% confidence intervals of: 8.1 to 11.7; 15.3 to 24.9; and 15.5 to 46.8. For p16 we previously had estimates of SDa, SDi and SDg of 19, 38 and 118; these values will be estimated with 95% confidence intervals of: 17.0 to 20.9; 32.7 to 42.8; and 82.2 to 154.3.

**Validation on prospective cohort**

***Methods***

Four to six 3-5micron sections will be cut from each of approximately 504 anonymised samples from the Head Neck 5000 cohort. Each case will be randomly assigned to one of up to 6 NHS laboratories. The laboratory will carry out the required tests (described above) for each of the 4 biomarkers (TILs, HPV-ISH, p16 and Survivin) on all the sections of that case. The results are then sent to InHANSE, and will be amalgamated with outcomes and baseline characteristic data of Head and Neck 5000. The amalgamated data will then be analysed by the University of Oxford.

***Sample size calculation***

High-risk patients treated with surgery and chemoradiotherapy (CRT) have a 3-year overall survival (OS) of 63% compared to those treated with CRT alone (42.5%). Therefore, to enable detection of a difference of 20% (HR 0.51) between surgical and non-surgically treated patients in high risk cohort, a sample size of 210 high-risk patients is required for 80% power and with $P < 0.05$ (see detailed analysis below). As high-risk patients constitute approximately 50% of patients with OPC, an overall sample size of 416 is required. To factor for missing data and missing or un-scorable biomarker samples, we have increased the overall sample size by 20% to 504.

***Planned statistical analyses***

It is important to ensure that inference methods are well calibrated and powerful, and require minimal reasonable assumptions to hold. We will undertake statistical tests to compare the survival probability for those patients whose treatment (surgery + CRT) coincides with the biomarker classifier (the Concordant group) against those patients whose assigned treatment (CRT) disagrees with the biomarker classifier (the Discordant

group). That is, consider those set of patients where clinicians have assigned to surgery-and-CRT. We can see if any of them would have been assigned solely to CRT by the biomarker classifier. We can then compare the outcomes of those patients who underwent surgery-and-CRT where the classifier assigned them to surgery-and-CRT versus those who underwent surgery-and-CDT where the classifier would have prescribed CRT alone. Under a null hypothesis that the classifier is no better than random we expect to see no difference in outcomes between the two groups. The same is true for those patients assigned to CRT where the classifier can be used to retrospectively identify the patients in this set that it would have switched treatment and assigned to surgery-and-CRT.

As a first step then, a hypothesis test could compare Concordant vs. Discordant survival times within each treatment group (CRT and surgery) separately. Fig 1 shows the effect size required for the test to be powered. The effect is measured in terms of the margin between the survival probability of Concordant and Discordant patients. The difference in the red lines shows the effect size of difference that we are powered for in those patients who received surgery-and-CRT; and the black lines show for those receiving solely CRT. The x-axis shows the concordance probability. At the point of greatest discordance, 50%, the sample size is largest as the classifier disagrees with the clinical assignment. At the extreme of 100% concordance we have no power as the classifier and clinical assignment are identical. For example, if the biomarker classifier agrees with the treatment assignment on 50% of the subjects – i.e. 0.5 on the horizontal axis -- the minimum difference in survival that could be detected by the test is 20%, namely 53% survival for the Discordant group vs. 73% for the Concordant group. These survival probabilities are anchored by the 3-year OS rate of 63% in high-risk patients treated with surgery +CRT.

Further to the formal tests we will also explore calibration of predicted survival probabilities in the Concordant group via AUC classifier analysis (C-index) and Hosmer-Lemeshow plots across each biomarker to identify mis-calibration and / or possible nonlinear effects.

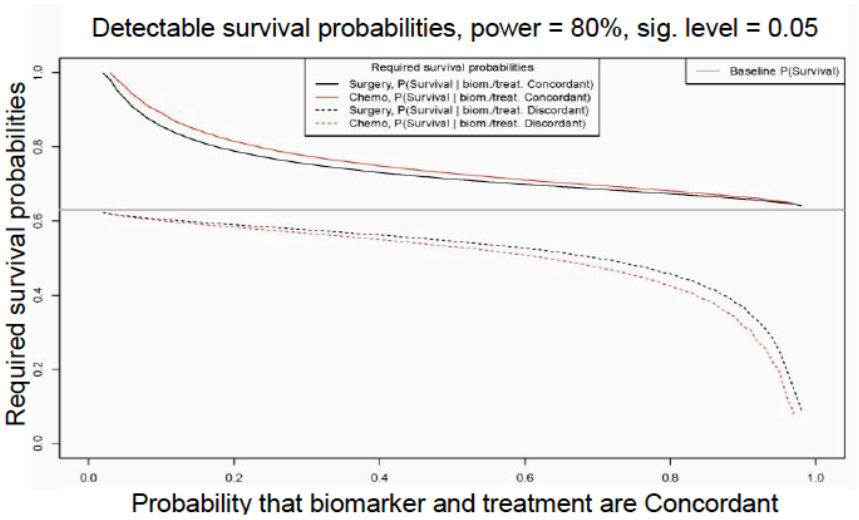


**Figure 1. Sample size calculation for project showing a detectable difference in survival probability between concordant and discordant groups at 80% power and 0.05 significance**

## Project oversight

Day-to-day project management will be undertaken by Professor Mehanna (CI) and Dr Rachel Spruce (Project Manager).

## References


1. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, et al. Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer. New England Journal of Medicine. Massachusetts Medical Society; 2010 Jul;363(1):24–35.

2. O'Sullivan B, Huang SH, Su J, Garden AS, Sturgis EM, Dahlstrom K, et al. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S): a multicentre cohort study. The Lancet Oncology. 2016 Apr;17(4):440–51.

3. Fakhry C, Zhang Q, Nguyen-Tân PF, Rosenthal DI, Weber RS, Lambert L, et al. Development and Validation of Nomograms Predictive of Overall and Progression-Free Survival in Patients With Oropharyngeal Cancer. Journal of Clinical Oncology. American Society of Clinical Oncology; 2017 Dec 20;35(36):4057–65.

4. Rainsbury JW, Ahmed W, Williams HK, Roberts S, Paleri V, Mehanna H. Prognostic biomarkers of survival in oropharyngeal squamous cell carcinoma: Systematic review and meta-analysis. Eisele DW, editor. Head & Neck. Wiley-Blackwell; 2012 Sep 20;35(7):1048–55.

5. Ang KK, Sturgis EM. Human Papillomavirus as a Marker of the Natural History and Response to Therapy of Head and Neck Squamous Cell Carcinoma. Seminars in Radiation Oncology. 2012 Apr;22(2):128–42.

6. Singhi AD, Westra WH. Comparison of human papillomavirus in situ hybridization and p16 immunohistochemistry in the detection of human papillomavirus-associated head and neck cancer based on a prospective clinical experience. Cancer. Wiley-Blackwell; 2010;92(9):NA–NA.

7. Schache AG, Liloglou T, Risk JM, Filia A, Jones TM, Sheard J, et al. Evaluation of Human Papilloma Virus Diagnostic Testing in Oropharyngeal Squamous Cell Carcinoma: Sensitivity, Specificity, and Prognostic Discrimination. Clinical Cancer Research. 2011 Oct 2;17(19):6262–71.

8. Marsh D, Suchak K, Moutasim KA, Vallath S, Hopper C, Jerjes W, et al. Stromal features are predictive of disease mortality in oral cancer patients. The Journal of Pathology. Wiley-Blackwell; 2011 Jan 5;223(4):470–81.

9. Ward MJ, Thirdborough SM, Mellows T, Riley C, Harris S, Suchak K, et al. Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. Br J Cancer. Nature Publishing Group; 2013 Oct 29;110(2):489–500.

10. Ambrosini G, Adida C, Altieri DC. A novel anti-apoptosis gene, survivin, expressed in cancer and lymphoma. Nature Medicine. Nature Publishing Group; 1997 Aug 1;3(8):917–21.

11. Xie S, Xu H, Shan X, Liu B, Wang K, Cai Z. Clinicopathological and Prognostic Significance of Survivin Expression in Patients with Oral Squamous Cell Carcinoma: Evidence from a Meta-Analysis. Wei Q-Y, editor. PLOS ONE. Public Library of Science; 2015 Feb 24;10(2):e0116517.

12. Jordan RC, Lingen MW, Perez-Ordonez B, He X, Pickard R, Koluder M, et al. Validation of Methods for Oropharyngeal Cancer HPV Status Determination in US Cooperative Group Trials. The American Journal of Surgical Pathology. 2012 Jul;36(7):945–54.

13. Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, et al. Prognostic Value of a Combined Estrogen Receptor, Progesterone Receptor, Ki-67, and Human Epidermal Growth Factor Receptor 2 Immunohistochemical Score and Comparison With the Genomic Health Recurrence Score in Early Breast Cancer. Journal of Clinical Oncology. American Society of Clinical Oncology; 2011 Nov 10;29(32):4273–8.