Author’s response to reviews

Title: Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population

Authors:

Mererid Evans (mererid.evans@wales.nhs.uk)
Robert Newcombe (newcombe@cf.ac.uk)
Alison Fiander (fianderan@cf.ac.uk)
James Powell (James.Powell@wales.nhs.uk)
Martin Rolles (martin rolles@wales.nhs.uk)
Selvam Thavaraj (selvam.thavaraj@kcl.ac.uk)
Max Robinson (max.robinson@newcastle.ac.uk)
Ned Powell (powelling@cf.ac.uk)

Version: 3 Date: 26 March 2013

Author’s response to reviews: see over
Dear Dr Chih-Yen Chien,

Re. “Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population.”

Thank you for taking time to consider this manuscript. We would also like to thank the reviewers for their thorough and helpful appraisal of this work. I have uploaded a revised manuscript, with changes tracked, which takes account of their comments. Point by point responses to comments are provided below (page numbers refer to the tracked changes version). I hope these are acceptable, and I look forward to hearing from you.

Yours sincerely,

Dr Ned Powell
Reviewer's report

Title: Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population

Version: 1 Date: 20 February 2013
Reviewer: Chung Feng Hwang

Reviewer's report:
The authors performed retrospective review the prevalence of HPV-positive OPC in an unselected UK population and correlate HPV positivity with clinical outcome. HPV was found in 55% of OPCs within the population. Dramatic improvements in loco-regional control and survival were seen in HPV-positives.

Major Compulsory Revisions

1. page 9 Statistical Methods: Analyses of OS and PFS included all patients, irrespective of treatment intent and response to treatment. -> The patient received palliative treatment should be excluded in the beginning because the clinical outcome was worse in this group.

This is a good point. We considered at length whether palliative patients should be included in the survival analyses, and after discussion with Prof Newcombe (Professor of Medical Statistics) we decided they should be included. The reasons for this decision were as follows:

1. The question we were seeking to address was, what is the difference in prognosis of HPV positive cases compared to HPV negative cases at a population level? Excluding a proportion of either group, based on stage at presentation and/or fitness for treatment, would introduce an unjustifiable bias. Not least because patients will be treated differently in different regions as the threshold for palliative management may vary between centres.

2. One of the strengths of the study was the inclusion of palliatively managed patients so that our cohort represents a “real world” population of patients with OPC and not a selected population. This provides a useful comparison with clinical trial cohorts, which tend to be a highly selected group.

3. We have been explicit about the composition of the groups and have also included summary data for survival analyses with the palliative patients excluded (pg14 para 2: “The effect of HPV status remained highly significant when palliative patients were excluded; OS at 3 and 5-years in radically treated HPV-positive patients was 82.6% and 75.4%, compared to 39.6% (95% CI: 32.5-46.7) and 31.1% (95% CI: 24.4-37.6) in HPV-negative patients, corresponding to a 74% reduction in the death rate (HR 0.259, 95% CI 0.152-0.440, p<0.001).”

2. page 17 2nd paragraph: This effect is greater than in many clinical trial cohorts, due in part (but not entirely) to the inclusion of palliative patients. -> Palliative patients should not be included in survival analysis.

This issue is partly addressed in the response above. The second paragraph on page 18 compares survival in a published trial cohort to survival in the study cohort. For this analysis, palliatively treated patients are excluded, and this is made clear: “Survival of radically treated HPV-positive patients was comparable to that reported in a large US study….”. This shows that palliative patients have been excluded when appropriate to the analysis.

3. page 19 conclusion: HPV was responsible for the development of 55% of OPCs in this study. -> There were no evidence to support HPV was responsible for the development of 55% of OPCs in the study. The authors only found HPV positivity in 55% OPCs cases.
The question of whether HPV is the underlying cause of a tumour is highly relevant. For HPV to be accepted as causally related to a tumour requires proof of expression of HPV encoded proteins in the relevant tumour tissue. There are two ways of doing this: the first is by demonstration of p16 over-expression; P16 is upregulated by release of the E2F transcription factor following E7 binding of Rb. The second is by detection of mRNA for the main HPV oncogenes (E6 and E7). In this study we have used presence of HPV DNA and evidence of p16 over-expression to define an HPV attributable tumour. This is an internationally accepted way to define an HPV driven tumour. This is well established and recognised within the field, and there are many publications to support this including:

Expression of p16 protein identifies a distinct entity of tonsillar carcinomas associated with human papillomavirus.

A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen
Smeets et al. Int. J. Cancer 2007; 121 2465

p16 expression as a surrogate marker for HPV-related oropharyngeal carcinoma: a guide for interpretative relevance and consistency

Singhi et al. Cancer. 2010 1;116(9):2166

4. page 19 conclusion: p16 IHC appears most prognostic and is unaffected by sample DNA quality, making it a useful test in clinical practice. -> The difference of p16 was 10% in the study. P16 was also affected by sample DNA quality. The sample size may be too small to reach statistical significance.

The effect of DNA degradation was assessed and is discussed on page 13 para 1: “DNA degradation did not have a significant effect on p16 IHC testing results. Estimated false negative rates are shown in Table 2.” It is correct that fewer p16 positives were observed among HMBS negative samples, but the association between p16 and DNA adequacy was not significant (p=0.28). A larger sample size might show a different result in either direction.

We accept that it would be helpful to stress that while p16 correlated best with prognosis, the difference between the three markers was not significant. This is already stated at the top of page 18: “The three HPV testing methods evaluated in this study were all good markers of survival, with no test performing significantly better than another.”

We have amended the following sentence: “In terms of point estimates, p16 performed best (HR for death 0.24), followed by ISH (0.27) and GP5+/6+ PCR (0.29), although the composite definition of HPV positivity (0.22) performed best overall.” to: “Hazard ratios (HR) for death were: 0.24 for p16, 0.27 for ISH, 0.29 for GP5+/6+ PCR and 0.22 for the composite definition of HPV positivity”.

We have also amended various sections to make this clearer. In the abstract “As a single marker, p16 was least affected by DNA quality and correlated best with prognosis”, has been amended to “As a single marker, p16 was least affected by sample quality and correlated well with prognosis”.

In the results (p12 para 3) “As a single marker, p16 performed best if all cases were included (HR for death 0.24, 95%CI 0.15-0.39), followed by ISH (0.27, 95%CI 0.16-0.46) and GP5+/6+ PCR (0.29, 95%CI 0.18-0.47), although all were slightly inferior to the composite definition of HPV-positivity (0.22, 95% CI 0.13-0.37),” has been amended to: “When all
cases were included, p16 correlated well with prognosis (point estimate of HR for death 0.24, 95%CI 0.15-0.39), as did ISH (0.27, 95%CI 0.16-0.46) and GP5+/6+ PCR (0.29, 95%CI 0.18-0.47), although all were slightly inferior to the composite definition of HPV-positivity (0.22, 95% CI 0.13-0.37), and no test performed significantly better than another."

In the conclusions (pg 20) “as a single marker, p16 IHC appears most prognostic and is unaffected by sample DNA quality, making it a useful test in clinical practice” has been amended to “p16 IHC appears prognostic...”

Minor Essential Revisions

1. page 6 line 2: in most other H&N cancers -> in most other H&N (head and neck) cancers
   “most other H&N cancers” has been amended to “most other Head and Neck (H&N) cancers”

2. page 19 1st paragraph: Although patients were identified retrospectively, their clinical and tumour data were recorded prospectively. -> The manuscript also belonged to a retrospective study.
   We agree that this wording is potentially confusing. The sentence has been deleted.

Discretionary Revisions

1. page 7 study population: This cohort represented ~50% of patients diagnosed with OPC during the period; the rest were not identified or excluded because pathology blocks were not retrievable. -> Why were -50% OPC pathological block not retrievable? Were the clinical data and outcome of the rest without pathological block different form the study population?
   We regret that the original “~50%” figure referred to cases diagnosed across the whole of Wales, as opposed to cases diagnosed within the study area (South Wales only). This has created a misleading and unduly negative picture of the dataset. Considering only the study area, there were 177 OPC diagnosed within the study period. Blocks were obtained for 147 of these i.e. 83% of possible samples were obtained. The relevant text has been amended from: “This cohort represented ~50% of patients diagnosed with OPC during the period.” to: “Histology blocks were obtained for 147 cases, representing 83% of patients diagnosed with OPC in South Wales during the period.”
   NB. This passage was previously at the start of the methods section, but has now been moved to the start of the results section.

   There are several reasons why blocks for the other 17% of cases were not obtained, including:

   1. In the interests of efficiency we focused collection of blocks on larger hospitals. A proportion of cases would have been biopsied at smaller hospitals from which we did not obtain material.
   2. There are likely to be some mismatches between coding in pathology databases and cancer registry databases i.e. some patients identified as OPC in registry databases were not classified as such in pathology databases.
   3. We obtained only a small number of cases from one major centre. This was largely due to logistical issues relating to the lack of an electronic patient database at the time the samples were collected.
   4. Some blocks were simply missing from, or could not be found in, pathology archives.
We have considered whether this would introduce significant bias into the data, and concluded that given the multi-factorial reasons for blocks not being included, this is unlikely.

To address this in the manuscript, the statement; “the rest were not identified or excluded because pathology blocks were not retrievable” has been deleted from the first paragraph of the Methods section. In the Discussion section the following passage has been added at the top of pg 20:

There are several potential limitations to the study. Histology blocks for 83% of OPC patients presenting across South Wales over the study period were obtained. There were several reasons why cases were not included: blocks were not collected from a number of smaller centres, there was limited collection from one major centre due to logistical difficulty in identifying the relevant cases, mismatches were observed in coding between registry and pathology databases, and some blocks were missing from pathology archives. There is no reason to suspect systematic bias in the sample, especially given the multi-factorial reasons for samples not being included, but the potential for some bias cannot be completely excluded.

2. page 9 p16 immunohistochemistry: p16 IHC was scored as positive if there was strong and diffuse nuclear and cytoplasmic staining present in greater than 70% of the malignant cells. -> The scoring system is subjective and not very popular.

p16 immunohistochemistry was carried out using the same antibody clone and scoring system presented in the Ang et al landmark publication in the New England Journal of Medicine (Ref: Ang KK et al. Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer, N Engl J Med. 2010). The same scoring system has been adopted by the Royal College of Pathologists (UK) in the ‘Dataset for histopathology reporting of mucosal malignancies of the pharynx’ (Ref: http://www.rcpath.org/publications-media/publications/datasets/pharynx) and is being used for oro-pharyngeal cancer clinical trials where p16 immunohistochemistry is incorporated into the trial registration (RTOG1016 in the USA and De-Escalate HPV in Europe). The scoring system is qualitative and relies on a pathologist to interpret the staining. In our study, two pathologists, with experience of interpreting the test, scored the sections independently and then came to a consensus on discordant scores. This information has been added to the material and methods section (p10 para1). The independent scores were concordant in 95% of cases (Kappa score 0.899). The same pathologists had reported similar concordance (97%) in a previous study (Ref: Thavaraj S et al. Evaluation of human papillomavirus testing for squamous cell carcinoma of the tonsil in clinical practice. Journal of Clinical Pathology. 2011;64:308-12).

The scoring system is being widely used in clinical practice and research and this method provides optimal assessment of the p16 immunohistochemistry.

3. page 11 HPV prevalence: Tumours were classified as HPV-positive if they contained HPV DNA (by GP5+/6+PCR and/or ISH) and overexpressed p16. -> This classification was not very popular. I think the prognosis in the HPV DNA positive group was also better than HPV DNA negative group in this study. The 'Equivocal' groups made the study more complicated.

This classification is one that has been used in several previous studies, as referenced in the text. Inclusion of equivocal cases is important as these have been found at high frequency in some studies (Jjunor et al. ref below), where it has been suggested that they may have different outcomes and treatment responses compared to ‘true’ HPV-positive and HPV-negative cases.

Benefit of chemotherapy as part of treatment for HPV DNA-positive but p16-negative squamous cell carcinoma of the oropharynx.
As discussed above, it is essential to provide evidence of both the presence of HPV genetic material and of expression of HPV encoded genes to classify a tumour as HPV-associated. It is also essential that the equivocal groups are defined in order to allow meaningful discussion of concordance between tests.

4. page 16 discussion 1st paragraph: P16 expression is not affected by DNA quality and may be utilized as a single marker of HPV infection in clinical practice, -> The p16 expression was also affected by DNA quality (from 57% to 47%). Does it mean we only measure p16 expression and not need measure HPV infection in the future study?
These issues are addressed in the text. The p value for association between p16 status and HMBS is 0.28 and thus is not significant, and in pg 12 para 1 we state: “DNA degradation did not have a significant effect on p16 IHC testing results.”

In relation to the second point, the Conclusions on page 20 include the following: “as a single marker, p16 IHC appears prognostic and is unaffected by sample DNA quality, making it a useful test in clinical practice. P16 as a single marker is not sufficient however for studies which aim to accurately report HPV prevalence, when p16 coupled to at least one test for HPV DNA (PCR/ISH) is recommended”. This is because some cases were equivocal i.e. despite being p16 positive, they were negative for HPV DNA and therefore cannot be classified (based on our results) as truly HPV positive in HPV prevalence studies.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
I declare that I have no competing interests
Responses to reviewer's comments.

Reviewer's report
Title: Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population
Version: 1 Date: 21 February 2013
Reviewer: Richard Shaw

Reviewer's report:
Overall this paper is of very high quality and contains important data which should be published.

It is important that all the data from this series is published (perhaps as a large supplementary table) to allow subsequent use in meta-analyses. Similarly it is important to clarify all details of lab techniques, specifications, PCR conditions, perhaps again as supplementary data - so that other groups can readily replicate the very high standards adopted by these authors.

We agree that the data should be available for inclusion in future meta-analyses. However our previous experience suggests that this is best arranged through direct interaction with researchers compiling such analyses, and we happy to make the data available for this purpose.

Regarding providing additional details of laboratory analyses, the ISH and p16 are highly standardised commercial assays and were conducted according to the manufacturer’s instructions additional detail is also provided in a previous publication from some of the authors which is referenced in the text on page 9(14). The GP5/6 PCR EIA and HMBS PCR are not commercial assays and we are happy to provide additional details relating to these assays. Additional files are included with the revised manuscript and these are referred to in the Methods section.

All of the comments below are relatively minor, with the exception of some of the statements made in the abstract, which at first glance may not all be particularly reflective of the data:

P3L2 - text gives impression that infection might improve prognosis in existing tumour (what they mean is that HPV mediated cases have better prognosis than non HPV mediated cases?)
(Similarly P19L8 “dramatic improvements”… no patient’s prognosis has been improved here – a what is meant that the two groups of patients have distinct clinical features and distinct prognosis)
(Similarly P3 L21 “HPV positivity alters the behavior of OPC” is a similar minor grammatical error)
We agree with these comments.
P3L2 now reads: “HPV-associated OPC appear to have better prognosis than HPV-negative OPC.” Instead of “HPV infection appears to be associated with improved prognosis”.

P20 para 2 now reads “Significantly better loco-regional control and survival were seen in HPV-positive cases.” instead of “Dramatic improvements in loco-regional control and survival were seen in HPV-positives”.

P3 L21 now reads: “HPV positive cases are clinically distinct from other OPC” instead of “HPV positivity alters the clinical behaviour of OPC”.

P7L3 – as around 50% of patients were not included in the data, it might be worth considering potential sources of systematic bias in discussion. Two possibilities spring to mind- firstly that patients having non-surgical treatment will have much less tissue available therefore less likely to meet DNA quality
requirements, secondly that patients who are HPV+ve tend to have smaller primary tumours than HPV-ve in many series – perhaps again less likely to have good DNA.

How have HPV+ve OPSCC patients with occult primary tumours been dealt with?

As described above, the proportion of blocks that were not obtained was 17%. We acknowledge that we cannot completely exclude some bias in our results, however, in response to the two specific points raised:

There was no evidence from our data that patients having non-surgical treatment were more likely to be HMBS negative. HMBS status was primarily influenced by the fixative used in different hospitals (as explained on p11), rather than on the amount of tissue available for analysis.

There was no significant difference in overall AJCC stage between HPV-positive and negative patients (Table 1). Furthermore, there were no significant differences in T stage between HPV-positive and HPV-negative cases in our series. HPV-positive cases had a HMBS positivity rate of 55.4% vs 39% for HPV-negative cases (table 1), thus there is no evidence from our data that HPV-positive cases were less likely to have good quality DNA.

The discussion (p20) has been amended to highlight the fact that we cannot completely exclude any bias from our results and now reads “There is no reason to suspect systematic bias in the sample, especially given the multi-factorial reasons for samples not being included, but the potential for some bias cannot be completely excluded.”

Patients with occult primaries were not included in this study as they would not have been classified as OPC.

P7L9 “local recurrence” seems to have been used as a term to include local, regional and locoregional recurrence – although this is clear in this paper, this may limit its usefulness in subsequent metanalysis and should be avoided.

We do not have data to separate local and regional recurrence and acknowledge that the terms ‘local’ and ‘locoregional’ recurrence had been used interchangeably in the text. The text has been amended so that recurrence either at the primary site or regional lymph nodes is now consistently referred to as ‘locoregional recurrence’ (p7 and p10).

P7L6 METHODS – there is no detailed account given about how clinical outcome data has been recorded, other than by database. How is a patient who dies in a peripheral hospital flagged at the database, how is the cause of death accurately determined? Other UK series use the ONS which has evident limitations but this group have a more reliable data source?

These details were not included in the text in the interests of space but every effort was made by the study team to ensure the validity of the clinical data presented. To reflect this, the text on p7 has been changed from:

“Data on clinicopathological characteristics and outcome was collected prospectively on an electronic database used at Velindre Cancer Centre, a regional referral centre serving a population of 1.5 million”.

To:

“Data on clinicopathological characteristics and outcome were obtained from an electronic health record used at the regional Cancer Centre. Deaths in peripheral hospitals were automatically fed into the electronic record. Where cause of death was not documented on the electronic record, it was elucidated by review of patient notes, review of clinic letters and/or discussion with General Practitioners. For every patient who was alive at the point of analysis but had not been seen in hospital for the preceding 12 months (eg had been
discharged from follow-up), the study team contacted the General Practitioner to ensure that the patient was still alive with no evidence of disease recurrence”.

P11L15- although the use of a DNA quality threshold is laudable, and doubtless an advance on the great majority of the literature, what is the rationale for using this amplicon, how many cycles of PCR, what PCR conditions etc. What was the effect of small variations in PCR machine, conditions and cycles of the overall informativity of the series? Is this test valid? 
The exact conditions for the HMBS PCR are now listed as an additional file. This amplicon was chosen as it is a well characterised primer set and is widely used PCR in laboratories performing HPV testing, hence it will be familiar to many researchers in the field (see refs below). In many ways the choice of amplicon for the control PCR is largely arbitrary; a highly efficient amplification of any appropriately sized region of human DNA would suffice.


P12L5 – P16IHC does not assess HPV prevalence, so this statement needs a grammatical adjustment
This passage has been amended from: “HPV prevalence was similar when analysed by p16” to “the proportion of positive samples was similar when analysed by p16”.

P12L17- “if all cases included” this is then presumably misleading data as the negative cases by DNA are false negatives. This data can be presented but this needs to be highlighted. This misinterpretation seems to be carried though to results and even abstract.
We clearly define HPV positive samples at the start of the Results section as “Tumours were classified as HPV-positive if they contained HPV DNA (by GP5+/6+ PCR and/or ISH) and overexpressed p16 [13, 14]”. This is the accepted and published definition and by using it, we ensure that our data can be compared with other studies, even when an assessment of DNA quality has not been made. In the section on P12 ‘Prognostic Value of HPV testing methods’ highlighted by the reviewer, the prognostic value of the individual testing methods is given in all cases and subsequently in HMBS cases only (P12L23).

P3L17- “as a single marker, P16 was least affected by DNA quality and correlated best with prognosis”. This seems the single most important correction to consider in this paper - as the message of this is likely to be taken from the published abstract without context / critical details that the authors present in the main text.
This statement seems to be incorrect at 4 levels and should be rethought, especially in the abstract:
1. how should a protein expression reasonably be expected to be affected by DNA quality? 
Is there any precedent for measuring degredation of a protein against a rather robust macromolecule such as DNA? – there is an assumption made here not explored in the detail.
2. although it is possible, it is evidently inappropriate, to compare P16 against DNA tests in samples that have proven inadequate DNA. The authors have moved this field on a little by establishing a minimum standard of DNA quality (although this in itself is not validated) and should restrict this analysis to that
cohort with adequate tissue.
3. There were no significant differences between the prognostic ability as all the C.I.s overlapped.
4. We know that other cases have p16 overexpression in absence of HPV aetiology (although the numbers are likely small) which seems to be ignored here.

We agree that this statement is liable to misinterpretation and have made the following amendments:

“As a single marker, p16 was least affected by DNA quality and correlated best with prognosis” has been amended to “As a single marker, p16 was least affected by sample quality and correlated well with prognosis”.

In answer to the specific points:
1. It is correct that DNA quality in and of itself will not influence measurement of protein abundance by IHC. This sentence was intended simply to convey that sample quality (assessed as DNA adequacy) did not have a significant impact on the p16 results.
2. We agree that it would not be appropriate to compare p16 vs DNA based tests in samples with inadequate DNA, but this is not what we have done. What we have done, is to evaluate markers that are already being used both in research and clinical practice, and determine the influence of adding in assessment of DNA quality. One of the clearest messages of the manuscript is that assessment of DNA quality is important, and that poor quality DNA can confound both PCR and ISH analyses. This is best illustrated by showing the impact of DNA quality assessment on a relevant cohort of material.
3. It is correct that the C.I.s overlapped. We have amended the text to make it absolutely clear that there were no significant differences between the tests:

In the abstract “As a single marker, p16 was least affected by DNA quality and correlated best with prognosis”, has been amended to “As a single marker, p16 was least affected by sample quality and correlated well with prognosis”.

In the results (p12 para 3) "As a single marker, p16 performed best if all cases were included (HR for death 0.24, 95%CI 0.15-0.39), followed by ISH (0.27, 95%CI 0.16-0.46) and GP5+/6+ PCR (0.29, 95%CI 0.18-0.47), although all were slightly inferior to the composite definition of HPV-positivity (0.22, 95% CI 0.13-0.37)," has been amended to: "When all cases were included, p16 correlated well with prognosis (point estimate of HR for death 0.24, 95%CI 0.15-0.39), as did ISH (0.27, 95%CI 0.16-0.46) and GP5+/6+ PCR (0.29, 95%CI 0.18-0.47), although all were slightly inferior to the composite definition of HPV-positivity (0.22, 95% CI 0.13-0.37), and no test performed significantly better than another.”

In the conclusions (pg 20) “as a single marker, p16 IHC appears most prognostic” has been amended to “p16 IHC appears prognostic...”

4. It is correct that most studies that use ISH and p16 do report some cases with upregulation of p16 without identification of HPV DNA. We have discussed this briefly in the second paragraph of the discussion: “The number of ‘equivocal’ cases with discrepant HPV DNA and p16 testing results was significantly lower than in some other studies [7, 18], suggesting that a testing algorithm combining PCR and ISH increases sensitivity for HPV DNA detection [14]. Discordant HPV DNA and p16 testing results occurred in 6-7% of cases showing that p16 alone is not sufficient for studies that aim to accurately report HPV prevalence”. We did not feel that we had data to support further speculation in this area.

P13L16 .. had a mean difference in age of..
“Patients treated with RT/CRT were 8.6 years older (mean age 61 vs 52 years) and had poorer performance status than those treated surgically” has been amended to:

P13L16 .. had a mean difference in age of..
“Patients treated with RT/CRT were older (mean age 61 years) and had poorer performance status than those treated surgically (mean age 52 years)”

P15L3 & P18L6. Further there are many clinical trials being designed around the potentially erroneous concept that smoking/HPV constitutes an entirely distinct prognostic entity. This data leans on survival estimates derived from evidently sub-optimal HPV “diagnoses” - this is a key point for discussion and possibly abstract as the clinical implications are important.

We agree that this is potentially an important observation that could have implications (if confirmed in other studies) for clinical trials and future clinical practice. We have amended P19 of the discussion from:

“Retrospective analyses have suggested that smoking can negatively affect survival in some HPV-positive patients [5, 7, 20]; the relatively small cohort (n=117) with known smoking history, crude definition of smoking used and/or large effect of HPV status on outcome may have masked the effect in this study. However, it is possible that the effect of smoking, particularly past smoking, on outcome from HPV-positive OPC has previously been over-estimated, an issue that should be addressed prospectively in future studies”

To:

“Retrospective analyses have suggested that smoking can negatively affect survival in some HPV-positive patients [5, 7, 20], and this data has influenced the design of several clinical trials. Although the relatively small cohort (n=117) with known smoking history, crude definition of smoking and/or large effect of HPV status on outcome may have masked the effect of smoking in this study, it is possible that the effect of smoking, particularly past smoking, on outcome from HPV-positive OPC has previously been over-estimated, and this issue must be addressed prospectively in future studies”

Although our data on smoking and HPV is interesting and hypothesis generating, it is not in our opinion sufficiently conclusive to include in the abstract.

P16L14 – It seems just as likely that the differences between HPV prevalence between the cited papers and this series might be due to dates of the samples analysed? (i.e. earlier samples will mean less HPV +ves, rather than a difference between the regions)

This is correct. The sentence has been amended from:

“It also adds to a picture of regional variation in HPV prevalence across the UK where rates of 37.5% (33/88) and 42.7% (77/180) have been reported [12, 18].”

to:

“It also adds to a picture of regional and temporal variation in HPV prevalence across the UK where rates of 37.5% (33/88) and 42.7% (77/180) have been reported [12, 18].”.

DISCUSSION – several papers point to RNA based tests being the gold standard – whether ISH or PCR – whilst discussing the relative merits of tests or combinations of tests, this might be worth a line or two in discussion?

We agree that several recent publications suggest that RNA based ISH may be useful in identifying HPV associated OPC. This is now acknowledged by amending:

“Although PCR-based testing protocols routinely incorporate assessment of DNA quality, ISH-based techniques do not and therefore risk under-estimating HPV prevalence; this may partly explain the lower sensitivity reported for ISH compared to other HPV detection methods in previous studies [17].”

To:
“Although PCR-based testing protocols routinely incorporate assessment of DNA quality, DNA-based ISH techniques do not, and therefore risk under-estimating HPV prevalence; this may partly explain the lower sensitivity reported for ISH compared to other HPV detection methods in previous studies [17]. However a recently developed RNA-based ISH test for HPV does include a control for sample quality and shows considerable promise as a diagnostic marker for OPC [19].

Level of interest: An article of outstanding merit and interest in its field
Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests:
I have published a peer reviewed research paper with one of the authors in the past (MR)
I am co-applicant with one of the authors on a clinical trial grant (ME)
Reviewer’s report

Title: Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population

Version: 1 Date: 17 February 2013

Reviewer: Tsung-Lin Yang

Reviewer’s report:

Major Compulsory Revisions

This study aims to study the significant impact of HPV infection on OPC. The authors states that HPV positivity alters the clinical behaviour of OPC, and is associated with significantly better clinical outcomes. Several different methods have been used to confirm the HPV positivity in the enrolled cases, and the correlation between HPV and clinical survival has also been clarified.

1. The analysis of HPV infection is mainly based on the FFPE which is retrieved from the pathological archive. Since the cases were enrolled from 2001 to 2006, many blocks were unable to be retrieved because of the quality of paraffin block. It may lead to a bias in the recruitment because the old blocks may be more difficult to be kept in good quality. 50% loss seems to be higher and may not be adequate for being presented as the whole population. A detail list and the clinical data of enrolled and excluded cases should be listed and compared to reveal the potential bias of block retrieval.

Please note that this comment is partly addressed in answer to a comment by Dr Chung Feng Hwang (the correct figure for the proportion of blocks included is 83%).

The suggestion that cases were excluded from the analysis based on the quality of the block is not correct. The reasons why blocks were not obtained included:

1. In the interests of efficiency we focused collection of blocks on larger hospitals. A proportion of cases would have been biopsied at smaller hospitals from which we did not obtain material.
2. There are likely to be some mismatches between coding in pathology databases and cancer registry databases i.e. some patients identified as OPC in registry databases were not classified as such in pathology databases due to poor coding.
3. We obtained only a small number of cases from one major centre. This was largely due to logistical issues relating to the lack of an electronic patient database at the time the samples were collected.
4. Some blocks were simply missing from, or could not be found in, pathology archives.

We regret that it is not possible to obtain clinical data for the patients who were not included in the study. We know of the existence of these patients from cancer registry data, but this does not identify individuals or record treatment data.

2. This study claims that the impact of HPV status on outcome is revealed in a ‘real-world’ population of patients with OPC without systematically excluding any patients. However, the selection bias may be encountered. The author needs to address in the discussion paragraph. Most palliative patients were noted in the true-negative group, which may change the analytic results of survival.

Survival was assessed, and is described both with and without inclusion of palliatively treated patients on page 14:
A clear association between HPV-positivity and favourable prognosis was demonstrated in Kaplan-Meier analysis (Fig 1A). 3 and 5-year OS rates were 82.6% (95% CI: 73.7 to 91.5) and 75.4% (95% CI: 65.2 to 85.5) respectively in HPV-positive patients, compared to 32.2% (95% CI: 20.3 to 44.1) and 25.3% (95% CI: 14.2 to 36.4) in HPV-negative patients, corresponding to a 78% reduction in death rate associated with HPV-positivity (HR 0.220, 95% CI; 0.132-0.366, p<0.001). Survival in patients with equivocal HPV status (n=10) was intermediate between that of ‘true’ HPV-positives and negatives (Fig 1B). The effect of HPV status remained highly significant when palliative patients were excluded; OS at 3 and 5-years in radically treated HPV-positive patients was 82.6% and 75.4%, compared to 39.6% (95% CI: 32.5-46.7) and 31.1% (95% CI: 24.4-37.6) in HPV-negative patients, corresponding to a 74% reduction in the death rate (HR 0.259, 95% CI 0.152-0.440, p<0.001).

3. In the true negative group, not all patients receive radical treatment. Similarly, not all patients receive radical management. No cases are noted in the true positive cases. This leads to a worse outcome in this group, and results in the analysis unconvincing. Our data shows that even when palliative cases are excluded, there is a significant difference in outcome between HPV-positive and negative cases (see comment in response to 2. above).

4. Since the quality of DNA is the most important factor for PCR analysis, how about the classification of 78 cases confirmed by HMBS positivity? how about the survival analysis based on the HMBS positive cases? It is also suggested to show the K-M plots based on these cases just like figure 1.

We are explicit that the survival analyses presented are based on the full 138 cohort so that the maximum amount of clinical data is presented. No significant differences in results were seen when HMBS positive cases only were analysed and these results were not presented in the interests of brevity and clarity. In view of the reviewers’ comments, we have included the following explanation on page 14:

“No significant difference in survival was seen when HMBS negative cases were excluded from the analysis and, as a result, HMBS positive and negative cases were combined for subsequent analyses, although every analysis was repeated in HMBS positive cases only to ensure that the results were consistent.”

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable
Reviewer’s report

Title: Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population

Version: 1 Date: 20 February 2013

Reviewer: MEIJIN NAKAYAMA

Reviewer’s report:
This is an article analyzing human papillomavirus-associated oropharyngeal cancer. The article is well designed and the analyses well conducted. The authors are recommended to comment in detail for the following points:

1) The contents of surgical interventions were not detailed. What kind of approaches was incorporated for the treatment, transoral or external approach? Did all patients who undergo surgery receive simultaneous neck dissections? Was there any correlation between different types of surgery and positivity of HPV?
These are very interesting questions but details of surgical technique were not collected as they were not relevant to the focus of the current manuscript. There are no obvious reasons to hypothesise a relationship between types of surgery and HPV status. Furthermore such an analysis would be very easily confounded by use of different surgical techniques at different centres, and changes in surgical practice over time.

2) The prognosis of HPV negative group was poor regardless of previous treatment, surgery or RT. Based on the current analyses, how the authors define the treatment selections, surgery or RT or CCRT, to the HPV positive and negative patients?
We are not entirely clear about the meaning of this comment. If the question is, can HPV status be used to guide choice of treatment, then based on the current analysis, the answer is no, as HPV positive patients had poor prognosis independent of treatment method and HPV negative patients had good prognosis independent of treatment method. (Pg 15 para 2 reads: HPV-positivity was also associated with better survival regardless of primary treatment modality (surgery or RT/CRT).

3) In order to reduce the incidence of distant metastases and second primaries, what do the authors suggest for the future management in this group of patients analyzed?
Our data does not allow us to make evidence – based suggestions on how to reduce the incidence of distant metastases and/or second primaries in patients with OPC. Further data on the frequency of occurrence and location of these, particularly second H&N primary malignancies, is required before recommendations regarding management can be made.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.