Article:
Sekhar, H, Zwahlen, M, Trelle, S et al. (8 more authors) (2017) Nodal stage migration and prognosis in anal cancer: a systematic review, meta-regression, and simulation study. The Lancet Oncology, 18 (10). pp. 1348-1359. ISSN 1470-2045

https://doi.org/10.1016/S1470-2045(17)30456-4

Nodal stage migration, the Will Rogers phenomenon and prognosis in anal cancer: systematic review, meta-regression and simulation studies

Hema Sekhar MRCS,[1] Marcel Zwahlen PhD,[2] Sven Trelle PhD,[3] Lee Malcomson BSc,[1] Rohit Kochhar FRCR,[4] Mark P Saunders PhD,[5] Matthew Sperrin PhD,[6] Professor Marcel van Herk PhD, [1] Professor David Sebag-Montefiore FRCR,[7] Professor Matthias Egger MD,[2] Professor Andrew G Renehan PhD.[1]

[1] Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[2] Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

[3] Clinical Trials Unit, University of Bern, Bern, Switzerland

[4] Department of Radiology, The Christie NHS Foundation Trust, Manchester, UK

[5] Department of Clinical Oncology, The Christie NHS Foundation Trust, Manchester, UK

[6] Farr Institute, MRC Health eResearch Centre (HeRC), Division of Informatics, Imaging and Data Sciences, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[7] Leeds Institute of Cancer & Pathology, University of Leeds, St James's University Hospital, Leeds, UK

Running title: Will Rogers phenomenon and prognosis

Keywords: Will Rogers phenomenon, prognostic discrimination, lymph node involvement, anal cancer

Correspondence to:

Professor Andrew G Renehan
Division of Cancer Sciences, School of Medical Sciences,
Faculty of Biology, Medicine and Health,
University of Manchester,
The Christie NHS Foundation Trust,
Wilmslow Road
Manchester,
M20 4BX UK
Tel: +44 161 3060870
E-mail: andrew.renehan@manchester.ac.uk

Abstract: 454 words; main text: 4862 words; 2 tables; 5 figures; 53 references ; supplemental material (56 pages); language: UK English.

ABSTRACT

Background: In patients with anal squamous cell carcinoma (ASCC), lymph node positivity (LNP) indicates poor prognosis for survival and is central to radiotherapy planning. Over the past three decades, LNP proportions have increased, mainly reflecting enhanced detection with newer imaging modalities; a process known as nodal stage migration. If accompanied by constant T stage distributions, prognosis for both LN+ and LN- categories may improve; a paradox termed the Will Rogers phenomenon. The latter has not been systematically evaluated, in general, or for anal cancer. Here, we aim to systematically evaluate the impact of nodal stage migration on survival in ASCC and address a novel hypothesis that this phenomenon additionally results in reduced prognostic discrimination.

Methods: We conducted a systematic review and meta-regression to quantify changes in LNP over time and the impact of this upon survival and prognostic discrimination. We searched Medline, Embase and the Cochrane Library (until 11 October 2016) to identify studies in patients with ASCC, where chemo-radiotherapy or radiotherapy was the main treatment, that (i) reported LNP proportions; and (ii) evaluated the relationship of lymph node status with prognosis. To investigate scenarios where reduced prognostic discrimination might occur, we simulated varying true LNP proportions and true survival rates, and compared these with expected observed outcomes for varying levels of misclassification of true nodal state.

Results: We included 62 datasets (10,569 patients). LNP proportions increased from a mean estimate of 15% (95% CIs: 10% to 20%) in 1980 to 37% (95% CIs: 34% to 41%) in 2012 ($p < 0.001$). In 11 studies with prognostic data, increasing LNP was associated with improved overall survival in both LN+ and LN- categories, while the proportions with tumour stage T3/T4 remained constant. In 20 studies, across a range of LNP proportions from 15% to 40%, the hazard ratios for survival of LN+ versus LN- decreased significantly ($p = 0.014$) from 2.5 (95% CI: 1.8 to 3.3) to 1.3 (95% CI: 1.2 to 1.9), demonstrating the phenomenon of reduced prognostic discrimination. The simulated scenarios reproduced this phenomenon where the true proportions for LNP were either 20% or 25%, but not where the true proportions for LNP were 30% or greater, arguing that the true proportion of LNP might be

lower than that observed in modern clinical series, which generally observe LNP proportions greater than 30%.

Interpretation: With nodal stage migration over time in anal cancer as an example, we describe a novel extension of the Will Rogers phenomenon, namely a form of misclassification that we have termed reduced prognostic discrimination. At a general level, the introduction of new staging technologies in oncology, which misclassify true disease stage, might spuriously inform management and ultimately have a risk of over-treatment.

Funding: Bowel Disease Research Foundation (BDRF)

Research in context

Evidence before this study

In oncology, the Will Rogers phenomenon occurs when patients are re-classified from one tumour stage to another stage following the introduction of either a new diagnostic technology or a new staging system, and there is a consequent paradoxical improvement in survival rates of both stages. A seminal article illustrating this phenomenon as a source of misleading survival statistics was published in 1985. Our search in Medline (01 Jan 1985 to 16 Oct 2016) identified only fourteen primary studies, which in the main, addressed confirmation of the Will Rogers phenomenon in 'before and after' analyses or comparisons of different staging systems, across various cancer types. We found no study that systematically evaluated the phenomenon.

Here, we extend the evaluation of the Will Rogers phenomenon with the example of anal cancer. For anal squamous cell carcinoma (ASCC), lymph node positivity (LNP) indicates poor prognosis for survival and is central to (chemo-)radiotherapy dose planning – the main modality of initial treatment. There are six published phase III trials of chemo-radiotherapy interventions in patients with ASCC. Secondary analyses from four of these show that, over the past three decades, LNP proportions have increased, mainly reflecting enhanced detection with newer imaging modalities. We hypothesise that this nodal stage migration (over time) is associated with improved prognosis in both LN+ and LN- categories, and if T stage at clinical presentation (the main prognostic factor for survival) otherwise remains constant, this fulfils criteria for the Will Rogers phenomenon. We additionally address a novel hypothesis that this occurrence results in reduced prognostic discrimination, as a form of misclassification, which in turn, might lead to over-treatment.

Added value of this study

Through systematic review, we used the strength of the literature of over ten thousand patients with anal cancer from 62 studies, and performed meta-regression demonstrating that, over the past three decades, there has been a near 7 percent increase in detection of lymph node positivity per 10 years. This nodal stage migration has occurred when the proportion of tumour stage T3/T4 remained relatively stable. By capturing this striking relationship between increased levels of LNP over time, we were able to infer that the factors driving the upward nodal stage migration (namely new imaging technologies) are concurrently: (i) improving prognosis for LN+ and LN- categories, thus fulfilling criteria for the Will Rogers phenomenon, while also (ii) resulting in a new observation, namely the

phenomenon of reduced prognostic discrimination. The simulated scenarios reproduced this phenomenon where the true proportions for LNP were either 20% or 25%, but not where the true proportions for LNP were 30% or greater, arguing that the true proportion of LNP might be lower than that observed in modern clinical series, which generally observe LNP proportions greater than 30%.

Implications of all the available evidence

The true performance characteristics (namely, sensitivity and specificity) of pre-treatment staging modalities are incompletely quantified for several cancer types in oncological practice because full histological confirmation of a carcinoma together with its lymphatic drainage field is generally absent. This case example of anal cancer (where pre-treatment staging is defined by imaging) illustrates the paradoxical possibility that the introduction of new and seemingly improved staging technologies in oncology might be associated with substantial misclassification of true stage; in turn, resulting in susceptibility to upward nodal stage migration, spuriously informing management, and ultimately a risk of over-treatment.

INTRODUCTION

Anal squamous cell carcinoma (ASCC) is a Human Papilloma Virus (HPV)-related malignancy[1] [2] whose incidence has increased substantially (2- to 4-fold) in both men and women over the past decades[3, 4]. Results from randomized trials, published in the 1990s, established combined chemo-radiotherapy (CRT) as the mainstay of initial treatment[5-7]; today about three-quarters of patients receive CRT as primary therapy[8].

Spread from the primary tumour is predominantly via lymphatic vessels to regional lymph node fields. Lymph node positivity (LNP) is an adverse prognostic factor, as reflected in the American Joint Committee on Cancer/International Union against Cancer (AJCC/UICC) staging system[9]. The 7th AJCC/UICC staging classification defined stages I and II by T1N0M0 and T2N0M0, respectively, while stage IIIA and IIIB are defined mainly by nodal positivity [T(1-3)N1M0; and T4N1M0, any TN2-3M0, respectively][9]. Detection of LNP is central for planning radiotherapy doses and fields. Thus, many centres traditionally use regimens based on the ACT II trial[10], with patients receiving radiotherapy doses of 50.4-54 Gy to both the primary tumour and the involved nodal fields (broadly equivalent to stage III tumours), with reduced doses (30.6-36.0 Gy) to uninvolved nodal draining fields (stage I and II).

Pre-treatment nodal staging is almost exclusively done by imaging. In the 1980s and 1990s, staging was by clinical examination, supplemented with computerised tomography (CT). In the 2000, magnetic resonance (MR) was introduced[2] and since 2010 Positron Emission Tomography (PET) has been recommended routinely or in selected patients to further enhance nodal staging[11-13]. Over the past three decades, LNP proportions have increased, probably driven by newer imaging technologies. For example, in trials performed in the 1990s, 17% of patients in RTOG 87-04[6] and 20% in ACT I[5] were lymph node positive, rising to 27% in RTOG 98-11(1998 to 2005)[14], and to 35% in ACT II[10], which recruited from 2001 to 2008, and where MR imaging was performed in approximately half of patients.

The observation of increased nodal positivity over time is known as nodal stage migration, and if accompanied by constant T stage distributions, there is potential for

paradoxical improvements in survival for both LN+ and LN- patients. This is known as the Will Rogers phenomenon[15]. The latter is a form of reclassification and is well recognised in the oncology literature after the introduction of new imaging modalities in other cancers (detailed references in webappendix p1-2), but has not been systematically evaluated, in general or for anal cancer. Here, we aim to systematically evaluate the impact of nodal stage migration on survival in ASCC and address a novel hypothesis that this phenomenon additionally results in reduced prognostic discrimination, with the hallmarks of narrowing survival differences between LN+ and LN- categories. If true, this would represent a form of misclassification and a risk for potential over-treatment.

METHODS

Study design

We conducted a systematic review and meta-regression to quantify changes in LNP over time and examined the associations between LNP and survival, and between LNP and prognostic discrimination. Here, we use the term prognostic discrimination as a clinical measure of a between category difference in survival rates, for example, at different time points or for different true LNP proportions. To determine whether nodal stage migration was accompanied by constant T stage distributions, in parallel, we quantified changes in T3/T4 proportions over time. Because development of imaging technology and expansion of multi-modality pre-treatment staging cannot be quantified directly, we used calendar time as a proxy measure. To better understand prognostic discrimination over time, we simulated hypothetical scenarios with varying levels of true LNP proportions; derived observed LNP proportions for varying test performance characteristics; and estimated the corresponding survival rates by observed nodal status. Throughout this paper, we use lymph node positivity (LNP) to indicate apparent clinical lymph node positivity determined by the study investigators at the time period for that study. For pragmatic reasons, we denoted LNP for involvement of any lymph nodal field draining the primary anal cancer (peri-rectal; iliac;

inguinal). The abbreviations LN+ and LN- were used to denote two strata of nodal status in prognostic modelling.

Search strategy, inclusion criteria and data extraction

We searched Medline, Embase and the Cochrane Library (until 11 October 2016) to identify potentially eligible randomised trials and observational studies in patients with ASCC. The search was limited to studies published in English from 01 Jan 1970 onwards, as RT or CRT was not used prior to 1970 (search details: webappendix p3-6. Reference lists of included papers were reviewed.

Five criteria had to be met: (i) unselected patients with anal cancer where treatment was primarily RT and/or CRT (including studies that employed a boost dose or phase 2 dose with brachytherapy or interstitial radiotherapy and studies with <10% treated with primary surgery); (ii) histological diagnosis of ASCC (<10% of other histologies were accepted); (iii) treatment with curative intent (<10% palliative or M1 disease were accepted); (iv) nodal status work-up that included either clinical examination, non-invasive imaging, or imaging and fine needle aspiration cytology (FNAC); (v) TNM stage or nodal status was reported. We excluded studies with fewer than 50 subjects for the following reasons: (i) risk of disproportionate influence in the random-effects meta-regression models; and (ii) many focused on reporting new technologies in highly selected patient groups.

The following data were extracted using a standardized, piloted form: study and participant characteristics; first and last years of enrolment, year of publication; T stage (7^{th} AJCC)[9]; method of nodal staging; and proportion of patients with LNP.

The primary outcome measure was overall survival (OS). For studies that reported survival data, we extracted 5-year OS rates for the total patient cohort, and by nodal status. Unadjusted and adjusted hazard ratios (HR) were extracted with standard errors (SEs) or 95% confidence intervals (CIs); for the main analysis, maximally adjusted HRs were used. If not reported, we calculated HRs from actuarial rates as $HR = \ln P_1 / \ln P_2$ (where P_1 is the survival rate of the LN+ group and P_2 the rate of the LN- group) and SEs as SE =

sqrt*(1/e1+1/e2) where e1 is the number of events in the LN+ group and e2 the number of events in the LN- group), or derived HRs based on data reconstructed from Kaplan-Meier survival curves for the whole study cohort and stratified by nodal status[16].

Assessment of risk of bias

We developed a five-component assessment tool to gauge the risk of bias, based on cohort selection; treatment type (selection of patient sub-groups for treatment other than chemo-radiotherapy); missing data; adjustment for potential confounding; and duration of follow-up. Not all criteria were relevant to each outcome. For example, follow-up was not assessed for studies of the proportion of patients with LNP. We classified risk of bias as high, moderate or low (detailed: webappendix p7). Studies where the risk of bias remained unclear were classified as high risk.

Statistical Analysis

For tables of study characteristics, summarised proportions across studies as mean proportions and ranges. For LNP proportions, we derived pooled estimates (and 95% confidence intervals, CIs) using the `metaprop` command in STATA.

We calculated study-level proportions and 95% confidence intervals of LNP patients and combined T3/T4 stages using exact binomial confidence intervals. Proportions were entered into random-effects meta-regression models[17] to assess their relationship with time. The majority of included studies were observational and there was an expected heterogeneity in the type and use of pre-treatment imaging for nodal staging over time for a given study. This heterogeneity was poorly reported. As an alternative, we sought to determine a time-point of 'average' clinical practice for the detection of LN+ over the period of a given study. We used median study year - that is year at which 50% of participants had been enrolled. We argued that for most treatment series, enrolment increases year on year. In the absence of parameters to estimate study-level medians, we used the 66th percentile year as an approximate based on in-house data (full justification: webappendix p8-10). In

sensitivity analyses, we repeated analyses using other timescales (first year of patient enrolment, last year of study enrolment, year of publication, and the 75th percentile year), and found similar patterns (webappendix 11).

Similarly, we used meta-regression to assess temporal trends of 5-year OS estimates by nodal status. We expressed changes in prognostic discrimination as changes in differences in survival rates, and examined associations between changes in prognostic discrimination and proportions of LN+ versus LN- patients in meta-regression models.

All meta-regression model were constructed using random-effects. We calculated prediction intervals and the between-study variance as τ^2.

For further sensitivity analyses, we sought to assess the potential impact of between-study heterogeneous characteristics on the main model and sequentially removed the following: studies with overlapping data (where data overlapped either in time or region without being duplicate data); studies that included patients with primary surgery, metastatic or palliative disease; and studies with histologies other than ASCC. We assessed for the influence of single studies on the summary estimates in a leave-one-out at a time approach. For prognostic studies, we addressed the effect of removing cohorts with fewer than 90% of patients treated with CRT, cohorts reporting univariate HRs only, and HRs estimated from actuarial rates and reconstructed data. Publication bias was assessed using funnel plots[18].

Simulations

In simulated studies, we explored mechanisms that may explain the effect of misclassified true node positivity on observed prognostic discrimination. We first created hypothetical single population of one million persons with varying levels of true LNP (20%, 25%, 30% and 35%). In a second step, we assumed that 5-year OS depends on true nodal status with 5-year OS set at 85% for true LN- and at 35% for true LN+ status (webappendix p12-15). With these assumptions, we simulated deaths assuming a Bernoulli distribution. In a third step, we derived a range of observed LNP proportions by varying the performance characteristics (sensitivity and specificity) of a hypothetical pre-treatment imaging test and then calculated

5-year OS according to observed nodal status. Finally, we regressed the average 5-year OS by the mean of observed LN status separately for positive and negative LN status. The mean LN status was obtained over different combinations of sensitivity and specificity. To define plausible ranges of sensitivity and specificity, we explored the literature for performance characteristics of pre-treatment imaging tests commonly used for anal cancer but found a limited number of studies; that meta-analyses were inconsistent [19, 20]; and concluded study quality was poor, mainly due to lack of a referent standard and high susceptibility to verification bias. Thus, we used a wide range of plausible sensitivities (5% increments from 40% to 95%) and specificities (1% increments from 68% to 97%), and varied both independently. Finally, to gauge which levels of true LNP were compatible with the observations from our systematic review, we plotted the mean 5-year OS by the mean of observed LNP.

All analyses were performed in Stata (version 14, Stata Corp., TX, USA) or in R, using R-3.2.2 (R Core Team, Austria, 2015).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

RESULTS

Included studies

The process of identifying eligible studies is shown in **Figure 1**. We included 62 studies that met all five eligibility criteria. Excluded studies and reasons for exclusion are detailed in webappendix p16-24. Forty-five studies contributed to analyses of 5-year OS; 11 studies to 5-year OS stratified by nodal status; and 20 studies contributed HRs comparing 5-year OS in LN+ and LN- patients.

Study characteristics

The 62 studies included 10,569 patients, and were published between 1984 and 2016; the year of last study ranged from 1979 to 2014. The pooled estimated LNP proportion at presentation was 28.8% (95% CIs: 26.5% to 31.1%) (webappendix p25-28). Data were mainly from institution-level treatment series but also included three randomised trials;[5, 6, 10] one national database;[21] and one cancer registry[22]. Ten pairs and one triplet of studies reported overlapping data (webappendix p29-30). The characteristics of the 45 studies (6,302 patients) and 11 studies (1,332 patients) included in the analysis of 5-year OS, and 5-year OS by nodal status, respectively, are detailed in webappendix p31-34. Median follow-up periods ranged from 1.2 to 8.6 years in the former group of studies, and from 3.4 to 8.2 years in the latter group. Pooled estimated LNP proportions were similar to that for the complete dataset of 62 studies: namely, 29.0% (95% CIs: 26.1% to 31.9%) and 27.0% (95% CIs: 21.6% to 32.3%). Estimates of 5-year OS were derived from reconstructed data from Kaplan-Meier curves in 17 and 3 studies, respectively. In most studies, CRT was either the sole or dominant treatment modality.

The majority of the 20 studies (4,048 patients) contributing HRs of 5-year OS comparing LN+ and LN- patients were published after 2000 (**Table 1**). Data were from one randomised trial[5]; one cancer registry[22]; and institution-level treatment series[23-40]. Median follow-up periods ranged from 1.3 to 8.6 years. Again, the pooled estimated LNP proportion was similar to that for the complete dataset of 62 studies: 27.7% (95% CIs: 24.2% to 31.2%).

Assessment of Study Methodological Quality

We assessed the methodological quality (webappendix p35) of all 62 studies reporting LNP proportions on two attributes: patient selection and missing data. The majority of studies were low risk. For the studies reporting 5-year OS for whole cohorts (45 studies) and by nodal status (11 studies), we assessed four attributes: cohort selection; treatment type; missing data; and follow-up. Many studies had moderate to high risk of bias of treatment type (due largely to the selection to use brachytherapy or interstitial therapy for radiotherapy

boosts); missing data (largely due to unclear reporting); and inadequate follow-up time (median follow-up < 5 years). For the 20 studies reporting HR, we assessed five attributes: cohort selection; treatment type; missing data; confounding; and follow-up. Again, many studies exhibited moderate to high risk of bias secondary to treatment type; missing data; and inadequate follow-up; and also confounding. Of the 20 studies, nine did not include adjustments. Where there were adjustments: all eleven studies adjusted for either T status, T size or N stage; four adjusted for histology; three adjusted for performance status; six and eight studies adjusted for age and gender, respectively.

Trends in lymph node positivity over time

In meta-regression analyses, there was a clear increase in observed LNP over time ($p<0.001$) **(Figure 2A)**. The proportion of apparent LN+ patients increased by 6.8% (95% CIs: 4.4% to 9.3%) per decade; the predicted mean LNP proportion was 15% (95% CIs: 10% to 20%) in 1980 increasing to 37% (95% CIs: 34% to 41%) in 2012. In contrast, the proportion with stage T3/4 disease remained essentially constant over time **(Figure 2B)**. Thus, we found evidence for nodal stage migration without corresponding change in T stage at presentation. There was little evidence for associations between LNP and other variables including study type, country, mean age, percent females, HIV status and tumour location (webappendix p36).

Survival by nodal status

We next examined whether nodal stage migration was associated with improved prognosis in LN+ and LN- patients (i.e., the Will Rogers' phenomenon). Based on 11 studies, the meta-regression analysis showed that there were indeed improvements in 5-year OS in both LN+ and LN- patients with increasing observed LNP **(Figure 3)**. For example, at 20% proportion of LNP, survival was 47% (95% CIs: 33% to 60%) in LN+ and 70% (95% CIs: 64% to 77%) in LN- patients, but 63% (95% CIs: 50% to 76%) and 78% (95% CIs: 71% to 85%), respectively, at a LNP of 35%.

For the analysis across 45 studies, consistent with the literature[8], 5-year OS rates increased from a mean estimate of 64% (95% CIs: 58% to 71%) in 1980 to 75% (95% CIs: 70% to 79%) in 2010 (p = 0.046]. Among ten of the 11 studies (with survival data by nodal status) and where full cohort survival was reported or derived, the 5-year OS rates increased over time in a similar manner to the model of 45 studies [64% (95% CIs: 51% to 76%) in 1980 to 72% (95% CIs: 63% to 81%) in 2010], though not statistically significant (webappendix p37).

Prognostic discrimination

For the 11 studies that reported paired OS rates, the prognostic discrimination between LN+ and LN- patients declined with increasing observed LNP. At a LNP proportion of 20%, the difference between mean predicted 5-year OS rates for LN- minus LN+ was 23% (70% minus 47%); at a LNP proportion of 35%, the difference was only 15% (78% minus 63%) **(Figure 3)**. We explored this further in a meta-regression of the 20 studies providing HRs for OS by nodal status. Across a range of LNP proportions from 15% to 40%, the HRs of LN+ versus LN- decreased significantly (p = 0.014) from 2.5 (95% CI: 1.8 to 3.3) to 1.3 (95% CI: 1.2 to 1.9) **(Figure 4)**. Similar patterns were noted with time as the x-axis (p = 0.052) (webappendix p38).

We investigated scenarios where reduced prognostic discrimination might occur using simulations and reproduced the findings from the meta-regressions of decreasing between nodal status survival differences within increasing true proportions of LNP. However, this phenomenon was limited to the scenarios of true LN+ proportions 20% and 25% (**Figure 5**), and absent in the scenarios of higher true proportions (30% and 35%). The patterns, and reductions in prognostic discrimination, for example, from 25.1 to 15.7 for true LNP 20% in the simulated scenarios mirrored those from the literature shown in Figure 3.

Table 2 shows the summary outputs from the simulations. For example, in the simulation of true LNP of 20%, observed node positivity of 20% arises where there is moderate sensitivity (mean: 63.5%) and high specificity (mean: 90.7%); however, a

moderate increase in sensitivity (mean: 68.5%) but without a proportionate maintenance of high specificity (decreased to mean: 73.3%) gives rise to a misclassification of observed node positivity of 35% (as reported in many contemporary studies). In the simulation of true node positivity of 35%, observed node positivity of 35% arise where there is moderate sensitivity (mean 66.5%) and high specificity (mean: 81.9%); a plausible scenario with modern imaging, but in the latter simulation, there is no reduced prognostic discrimination. The presence of reduced prognostic discrimination in the scenarios of true LNP of 20% and 25% argue that the true proportion of LNP might be lower than that observed in modern series (typically 35%).

Sensitivity analyses

For the meta-regression of LNP proportions over time (62 studies), similar results were found after excluding: studies with overlapping data; studies with patients treated with primary surgery; studies with patients with M1 disease; and studies that only reported inguinal node detection (webappendix p39). On the modelling of hazard ratios for overall survival by nodal status (20 studies), we repeated the meta-analyses stratified by study characteristics and designs, and tested for interactions, and found none (webappendix p40). We found similar summary risk estimates obtained from: studies where over 90% of patients received CRT versus those where less than 90% of patients received CRT; unadjusted versus adjusted HRs, both for all studies and where same studies reported both; or how HRs were derived (directly from Cox model, survival rates, or reconstructed data).

We noted that some studies, for example, that of Salmon et al.[23] with a last year of patient enrolment at 1979 might be an outlier. We repeated the meta-regressions of LNP versus median study year, excluding one study at a time, and found similar results (webappendix p41).

Publication Bias

A funnel plot of the twenty studies with HR of OS stratified by nodal status was constructed to assess for publication bias (webappendix p42). There was no clear evidence of asymmetry.

DISCUSSION

Summary of main findings

Among patients with anal cancer, our meta-regression demonstrated that, over the past three decades, there has been a near 7 percent increase in detection of lymph node positivity per 10 years. This nodal stage migration has occurred when the proportion of stage T3/T4 disease remained relatively stable, and is likely to be driven by increased sensitivities of new imaging technologies. The increase in LNP proportions was associated with improved 5-year OS rates in both LN+ and LN- groups of patients, thus fulfilling criteria for the Will Rogers' phenomenon. We observed a new phenomenon, a reduction in prognostic discrimination, with a reduction in study-level hazard ratios for nodal positivity over time. This is a misclassification and represents an extension of the process of stage migration. The simulated models replicated the reduced prognostic discrimination, but only for true LNP proportions of 20% and 25%, arguing that the true proportion of LNP might be lower than that observed in modern clinical series, which generally observe LNP proportions greater than 30%. At a general level, the introduction of new pre-treatment staging technologies and biomarker in oncology might be associated with substantial misclassification of true stage; in turn, spuriously informing management with risk of over-treatment.

Context of other literature

To our knowledge, this is the first large-scale review with meta-regression, in patients with anal cancer, of the prevalence of lymph node positivity; temporal changes in proportions with LNP; and the impact of these on prognosis. Lymph node positivity is considered a major adverse prognostic factor in anal cancer, as reflected in the AJCC/UICC staging system [9] and in recent prognostic studies in trial patients [41]. Its importance in treatment algorithms is

reflected in current guidelines in North America [42] and Europe [13]. Researchers have explored opportunities to refine the definitions of pre-treatment nodal status – for example with sentinel node biopsy [43, 44]. However, with increasing reliance on PET-CT for the detection of lymph node positivity [42], sentinel node biopsy has not gained widespread clinical use in patients with anal cancer. As initial radiotherapy treatment in patients with anal cancer moves away from conventional to a multi-field (modulated) technique [42], identification of nodal involvement will remain important.

The Will Rogers phenomenon is well-known in the cancer literature. This typically occurs following the introduction of new diagnostic or staging tests, which generally has increased sensitivity [15, 45, 46], thus allowing the detection of metastatic tumour deposits that were previously occult. The phenomenon may also occur after the introduction of new staging criteria[47]. This results in the reclassification of more patients into a higher disease stage or prognostic score, apparent stage-specific improvements in prognosis, but without altering survival in the individual patient – hence, the paradox. As a secondary finding, we showed improvement in overall survival with time, as noted in a UK population-based study.[8] In their classical paper in 1985, Feinstein et al. [15] indicated that "the total survival rate in the cohort (is) unaffected", which is true over a short time span. But in the present study (and indeed in the lung cancer example in the classical paper [15]), over a longer duration, there may be gradual improvement in survival due to period effects, such as treatment centralisation, improvements in medical management and in management of treatment-related morbidities.

To date, the implication of the Will Rogers' phenomenon has been mainly to caution researchers of the limits of using historical control groups for comparison on outcomes. Here, we demonstrate the additional phenomenon of reduced prognostic discrimination. This has wider clinical implications: (i) risk of over-diagnosis and over-reliance on nodal stage as a prognostic marker to guide treatment; and (ii) risk of under-powering a trial, if for example, a new diagnostic modality is introduced into a trial in addition to a new intervention.

It is important to note that this study did not aim to establish that newer imaging modalities were the cause of the observed Will Rogers' phenomenon, but rather to determine the impact of nodal stage migration on survival. Indeed, influences other than improved imaging technologies may have contributed, such as treatment centralisation and the development of anal cancer multidisciplinary meetings. Additionally, there might be biological explanations. For example, in oropharyngeal carcinoma, the proportion that is HPV-positive has increased with time, and in turn (compared with non-HPV-related), these HPV-positive tumours are associated with higher lymph nodal positivity; yet are more radio-sensitive, and hence, have a better oncological outcome[48, 49]. This mechanism might operate for ASCC, though less likely to substantially influence survival as near 95% of ASCCs are HPV positive[1].

Limitations and strengths

Our study has limitations. First, meta-analyses of observational studies are prone to the biases and confounding factors that are inherent in the original studies[50]. To address this, we assessed the methodological quality of key study-level domains, and found that many studies used in the prognostic analyses were at moderate to high-risk for inadequate adjustment for confounding. We compared, where possible, summary estimates for studies that reported univariate and multivariable analyses, and found no material differences. Second, we assumed that the period effect is, in the main, driven by introduction of new imaging technologies. However, data on imaging were difficult to extract from studies as these frequently were not well-described and often included a heterogeneity of imaging modalities. As a pragmatic alternative, we took study-level median study year as a surrogate of the representative imaging modality used in a given study. Third, reporting of survival data was frequently incomplete or missing. To counter this, we reconstructed data from Kaplan-Meier curves, based on a validated algorithm [16]. We made assumptions to derive hazard ratios in studies that reported survival rates only – to mitigate against this by undertaking meta-analyses with and without these studies, and found no material differences. We still

failed to capture all reported survival stratified by nodal status, such that there might be a differential outcome reporting bias. Fourth, there was considerable unexplained between-study heterogeneity in most of meta-regression models.

Additionally, there are temporal changes in treatment-related factors that are sources of potential residual confounding. First, there is known wide variation in radiotherapy practice, and in particular, whether prophylactic inguinal nodal radiation is used. Different treatment modalities might influence loco-regional failure rates, but data from the six published randomized trials in this field suggest that the latter does not influence survival (webappendix p43). Second, technical refinements in salvage surgery might account for some increase in long-term disease-free states, but primary loco-regional failure rates have reduced substantially (from 35% in the 1990s[5] to approximately 20% by 2010) such that salvage surgery is now less often required. Furthermore, among patients with local relapses, the proportions that proceed to salvage surgery has decreased from over 70% in historic series[25, 51] in the 1990s to only 23% in the ACT II trial, from the mid-2000s.[52] Third, enhanced detection and treatment of distant metastases might improve survival, but metastatic disease is relatively uncommon in anal cancer and most phase II chemotherapy trials show no significant benefit in this setting (reviewed elsewhere[2]). Finally, advances in the management of HIV-patients may account for some survival improvements. But, while HIV-positivity is a risk factor for incident ASCC, with the exception of five highly-selected US/German/French institute series, HIV positive cases generally account for less than 11% of anal cancers (webappendix p25-28) and improvements in HIV-related treatment is unlikely to substantially alter survival rates in the totality of anal cancer patients.

The study has a number of strengths. First, we exploited the power of meta-analysing many studies, with moderately large numbers of patients with anal cancer (an uncommon cancer), including data published over three decades. This temporal 'library' of data would not be very common for a single institute series. Second, we harnessed the heterogeneity of the meta-analysis of LNP proportion over time to demonstrate a clear temporal relationship and speculate that one explanation is through the different modalities of imaging. We

recognise that, in modern clinical practice, imaging modalities are commonly used in combination – our approach may have simplified this, but was pragmatic. Third, to test the clinical impact of the increasing rate of LNP with time, we additionally explored (through meta-analyses) the relationships with LNP proportion and survival, and through this, for the first time, we described the reduction of prognostic discrimination, as an extension of the process of stage migration.

Clinical implications and future research

Identification of lymph node positivity at the pre-treatment stage is central to chemo-radiotherapy planning. Modern clinical practice in many European centres is with IMRT (Intensity Modulated RadioTherapy), but the principles remain - reduced prophylactic inguinal irradiation in node negative patients; and a gross tumour dose of 50.4 – 54.0Gy in 25-28 fractions to areas of primary tumour and lymph nodal involvement. The latter is associated with greater early and late treatment-related morbidity, compared with the former. With increasing LNP proportions, the proportion of patients requiring higher RT doses and volumes increases and may translate into higher treatment-related morbidity, but with continually changing management, this hypothesis is near impossible to test.

The concept of stratified treatment in anal cancer is being explored with the current availability of treatment intensification using IMRT – for example, in the UK PLATO trial[53]. This approach aims to provide dose intensification to those patients thought to be at high-risk of disease relapse (including node-positivity) and provide more conservative radiotherapy doses to those at low-risk, to reduce toxicity profiles. Our study highlights the importance of continually evaluating the modalities that classify nodal status in patients with anal cancer, to inform future trials of stratified treatment approaches.

Conflict of Interest

AGR has received lecture honoraria from Merck Serona and Janssen-Cilag, and independent research funding and lecture honoraria from Novo Nordisk and MPS Sanofi Pasteur. MZ have received grant support from the European Union, Swiss National Science Foundation and from the Swiss Cancer League, which paid for PhD and post-doctoral positions. MZ has consulted for BIOTRONIK AG and NOVIGENIX, Switzerland, and has or had projects that received unrestricted research support from the following: AstraZeneca AG, Aptalis Pharma Inc., Dr Falk Pharma GmbH, Glaxo-Smith-Kline, Nestlé S.A., Regeneron Inc., and the Swiss Medical Association. All other authors have no conflicts of interest to declare.

Contributions

HS performed literature searches, data extraction, and contributed to analyses, data interpretation and writing of the manuscript. MZ and ME contributed to the design of the study, data analysis and interpretation, and writing. ST contributed to reconstruction of data from Kaplan-Meier curves. LM assisted with the literature screening and data extraction. MS contributed to statistical interpretation. RK, MPS, MvH and DS-M contributed on clinical and radiological interpretation. AGR conceptualized the paper and contributed to all sections of the manuscript.

Acknowledgements

We acknowledge the generous funding from Bowel Disease Research Fund.

Table 1: Twenty studies included in analysis of prognostic influence of nodal status (all had hazard ratios for overall survival)

Author	Study years	Study design	n	Node positive (n, (%))	Treatment	N (%) treated with CRT	Median follow up (range)	Confounders adjusted for	5-yr actuarial rates (LN- vs LN+) (95% CI)	Hazard Ratio (95% CI)	Estimated from actuarial results or reported HR	Actuarial OS if used reported in paper or obtained from KM curve reconstruction
Salmon – 1984 France[23]	1968 - 1979	Institute treatment series	183	23 (12.6)	RT	0	Not stated	No adjustments		3.23 (1.91 – 5.47)	Estimated	Reconstructed
Svensson - 1992, Sweden[24]	1985 - 1990	Institute treatment series	82	12 (14.6)	RT +/- Bleomycin	43 (57)	2.7 (0.08– 7.09)	Age, T stage, nodal status, gender, tumour site, histology		1.87 (0.62 -5.61)	HR - MVA	Not available
Arnott – 1996 UK[5] Gynne-Jones – 2013 UK[41]	1987 - 1994	RCT	275 CRT only	60 (21.8)	RT vs RT+5FU+MMC (prognostic analysis only uses CRT arm)	275 (100)	3.5 (IQR: 2.3 – 5.2)	Gender, nodal status, age, White blood cell count, Haemoglobin		1.74 (1.17 – 2.58)	HR - MVA	Not available
Grabenbauer – 2005 Germany[26]	1985 – 2001	Institute treatment series	101	27 (26.7)	RT + 5FU + MMC	101 (100)	7.5 (1 – 16)	No adjustments		1.74 (0.81 – 3.76)	Estimated	Reported
Nilsson – 2005, Sweden[25]	1985 - 2000	Institute treatment series	308	81 (26.3)	RT +/- bleomycin/Carboplatin/Cis	156 (51)	5.5 (1 – 16)	Gender, age, tumour site, histology, tumour size, nodal status		1.50 (0.95 – 2.37)	HR - MVA	
Das – 2007 USA[27]	1992 – 2004	Institute treatment series	167	65 (38.9)	RT + 5FU + Cis/MMC	167 (100)	3.5 (0.2– 12.2)	nodal status, HIV status		1.54 (1.18 – 2.00)	HR - MVA	Not available
Widder – 2008 Austria[28]	1990 – 2002	Institute treatment series	129	31 (24.0)	RT + 5FU + MMC	95 (74)	3.9	T stage, Nodal status, age, gender, overall treatment time, split course, total radiation dose, RT and chemotherapy		2.09 (1.21 – 3.62)	HR - MVA	
Fraunholz – 2011 Germany[29]	1997 - 2008	Institute treatment series	70	19 (27.1)	RT + 5FU + MMC	80 (100)	4.25 (0.3– 19.6)	No adjustments		2.83 (1.10 – 7.27)	Estimated	Reported
Wolfff – 2011 Germany[30]	1992 – 2004	Institute treatment series	72	22 (30.6)	RT + 5FU + MMC	65 (90.3)	7.7 (0.8– 17.2)	No adjustments		3.02 (1.37 – 6.70)	Estimated	Reported
De Foe – 2012 USA[31]	2003 - 2009	Institute treatment series	78	25 (32.1)		77 (98.7)	1.3 (0 – 6)	No adjustments		0.76 (0.28 – 2.10)	Estimated	Reconstructed
Chuong – 2013 USA[33]	2000 - 2011	Institute treatment series	89	29 (32.6)	RT + 5FU + MMC	89 (100)	2.2 (0.3– 11.1)	Treatment, age, gender, T stage, nodal status, RT break		2.02 (0.40 – 10.31)	HR - MVA	Not available
Eng – 2013 USA[35]	1989 - 2009	Institute treatment series	201	76 (37.8)	RT + 5FU/capcitabine + Cis	100	8.6	Not stated		1.15 (0.88 – 1.51)	HR - UVA	Not available
Fakhrian – 2013 Germany[36]	1988 - 2011	Institute treatment series	138	76 (29.0)	RT + 5FU + MMC	116 (84)	8.6	T Stage, UICC clinical stage, histology, ECOG performance status, radiotherapy technique, gender		1.28 (0.62 – 2.64)	HR - MVA	
Kim – 2013 Korea[34]	1979 - 2008	Institute treatment series	50	18 (36.0)	RT + 5FU + MMC	49 (98)	5 (0.7– 16.8)	Gender, performance status, nodal status, clinical complete		2.80 (0.60 – 12.07)	HR - MVA	Not available

								remission.				
Lestrade – 2013 Italy[32]	1993 - 2009	Institute treatment series	76	14 (18.4)	RT + 5FU + MMC	39 (51)	5 (0.6– 17.0)	No adjustments	75.9% (65.2 – 86.6) vs 75.7 (53.2 – 98.2)%	1.01 (0.31 – 3.28)	Estimated	Reported
Toh – 2013 UK[37]	2004 - 2011	Institute treatment series	92	30 (32.6)	CRT (no further details provided)	92 (100)	2.7 (1.3 – 5.0)	Gender, neutrophil:lymphocyte ratio		1.30 (0.39 – 4.35)	HR - MVA	Not available
De Bari – 2014 Switzerland[39]	2010 – 2012	Institute treatment series	100	29 (29.0)	RT + 5FU + Cis	58 (58)	3.4 (0.1– 19.8)	No adjustments	62% (50.7 – 73.3) vs 59% (41.4 – 76.9)	1.10 (0.56 – 2.18)	Estimated	Reported
Leon – 2014 Sweden, Norway, Denmark[22]	2000 - 2007	Registry based study	1266	402 (31.8)	RT + 5FU +/ MMC/5FU + Cis	669 (53)	4.2 (1.0 – 9.1)	Age, gender, TNM stage, performance status, non-protocol treatment		1.58 (1.25 – 2.18)	HR - MVA	Not available
Yeung – 2014 USA[38]	2000 - 2010	Institute treatment series	169	44 (26.0)	RT + 5FU + MMC	169 (100)	3.3	smoker, ECOG performance status, squamous cell histology, N status, T stage, MMC2 chemo		3.39 (1.54 – 7.49)	HR - MVA	Not available
Oblak – 2016 Slovenia[40]	2003 - 2013	Institute treatment series	100	35 (35.0)	RT + 5FU + MMC	100 (100)	4.3 (0.1– 10.8)	No adjustments		2.06 (1.09 – 3.90)	Estimated	Reported
		Total Pooled proportion estimate (95% CIs)*	**3740**	**27.7 (24.2 – 31.2)**								

HR = Hazard Ratio. CRT = chemoradiotherapy. RT = radiotherapy. 5FU = 5 – fluorouracil. MMC = Mitomycin C. Cis = Cisplatin
UVA = Univariate. MVA = Multivariable. KM = Kaplan-Meier.
IQR: inter-quartile range. CI: confidence interval.
*derived using the `metaprop` command in Stata

Table 2 The effect of misclassified true lymph node positivity (LNP) on observed prognostic discrimination

	Lymph node positivity (LNP) classification							
	True LNP 20%		True LNP 25%		True LNP 30%		True LNP 35%	
Mean observed LNP	20%	35%	20%	35%	20%	35%	20%	35%
Mean (range) sensitivities for mean observed LNP* %	63.5 (40-90)	68.5 (45-90)	55.8 (40-75)	67.0 (40-90)	49.3 (40-60)	67.6 (40-90)	45.0 (40-50)	66.5 (40-90)
Mean (range) specificities for mean observed LNP* %	90.7 (84-97)	73.3 (68-79)	91.9 (92-97)	75.6 (68-84)	92.6 (88-97)	78.9 (68-89)	93.3 (90-97)	81.8 (68-96)
	5-year survival rates %							
Observed LN negative mean survival rate (95% CIs) †	79.8 (79.6-80.1)	80.7 (80.3-81.0)	77.9 (77.4-78.3)	79.5 (79.1-79.8)	75.4 (74.9-76.0)	78.0 (77.7-78.3)	72.5 (71.8-73.2)	76.1 (75.8-76.5)
Observed LN positive mean survival rate (95% CIs) †	54.8 (53.9—55.6)	64.9 (64.0-65.8)	51.0 (50.0-52.1)	59.4 (58.6-60.2)	48.5 (47.3-49.7)	54.9 (54.1-55.6)	47.0 (45.6-48.3)	51.4 (50.8-52.1)
Mean survival (%) difference (LN- versus LN+) at mean observed LNP of 20%	25.1 (23.9-26.2)		26.8 (25.4-28.3)		26.9 (25.1-28.7)		25.6 (23.6-27.6)	
Mean survival (%) difference (LN- versus LN+) at mean observed LNP of 35%		15.7 (14.5-16.9)		20.0 (18.9-21.2)		23.1 (22.0-24.1)		24.7 (23.7-25.7)
Reduction (%) in prognostic discrimination at mean observed LNP 20% versus 35%	9.3 (9.2-9.4)		6.8 (6.5-7.1)		3.8 (3.1-4.5)		0.9 (-0.1-1.9)	

* Based on distributions of sensitivities and specificities for observed LNPs between 19% and 21%, and between 34% and 36%, respectively, for assigned mean observed LNP 20% and 35%.

† Based on regression models for the assigned mean observed LNP 20% and 35%.

FIGURE LEGENDS

Figure 1 1 Flow diagram of articles identified and screened for eligibility

[a] 23 datasets contributed from K-M reconstructions

[b] 5 datasets contributed from K-M reconstructions

[c] 2 datasets contributed from K-M reconstructions

Figure 2 A. Meta-regression of reported lymph node positivity (LNP) versus median study year. Circles represent individual studies (62 studies). **B**. Meta-regression of proportion of T3/T4 tumours versus median study year. Circles represent individual studies (61 studies). For A and B, the sizes of the circles are proportional to study sample sizes (ranging from 50 to 1266 participants). Dashed lines are 95% confidence intervals (CI) and prediction intervals (PI).

Figure 3 Meta-regression plots of 5-year overall survival against LNP stratified by nodal status. The vertical dashed grey lines indicate paired LN+ and LN- categories per study. The vertical double arrowed lines are mean differences in survival at 20% and 35% LNP. The values in parentheses are modelled 5-year overall survival rates at the respective LNP. The sizes of the circles are proportional to the sample sizes by nodal status.

Figure 4 Meta-regressions of LNP against the hazard ratio of the effect of nodal status on overall survival. Dashed lines are 95% confidence intervals (CI) (red) and prediction intervals (PI) (yellow). Circles represent individual studies (20 studies); the sizes of the circles are proportional to study sample sizes (ranging from 50 to 1266 participants). Note: the y-axis is log scale.

Figure 5 We constructed simulated scenarios with varying true LNP proportions (20%, 25%, 30% and 35%). The true survival rates were fixed at 85% and 35%, respectively. Survival

rates for the observed LN- versus LN+ states are plotted. The prognostic discrimination at 20% and 35% observed node positive proportions for each scenario are summarised in Table 2. Shaded areas are 95% confidence intervals (derived from the standard error of the forecast, and thus, are equivalent to predictive intervals in Figure 3); continuous bold lines are mean survivals.

References

1. Baricevic I, He X, Chakrabarty B, et al. High-sensitivity human papilloma virus genotyping reveals near universal positivity in anal squamous cell carcinoma: Different implications for vaccine prevention and prognosis. European Journal of Cancer 2015; **51**(6): 776-85.
2. Glynne-Jones R, Renehan A. Current treatment of anal squamous cell carcinoma. Hematol Oncol Clin North Am 2012; **26**(6): 1315-50.
3. Islami F, Ferlay J, Lortet-Tieulent J, Bray F, Jemal A. International trends in anal cancer incidence rates. Int J Epidemiol 2016.
4. Wilkinson JR, Morris EJ, Downing A, et al. The rising incidence of anal cancer in England 1990-2010: a population-based study. Colorectal Dis 2014; **16**(7): O234-9.
5. Arnott SJ, Cunningham D, Gallagher J, et al. Epidermoid anal cancer: Results from the UKCCCR randomised trial of radiotherapy alone versus radiotherapy, 5-fluorouracil, and mitomycin. Lancet 1996; **348**(9034): 1049-54.
6. Flam M, John M, Pajak TF, et al. Role of mitomycin in combination with fluorouracil and radiotherapy, and of salvage chemoradiation in the definitive nonsurgical treatment of epidermoid carcinoma of the anal canal: results of a phase III randomized intergroup study. J Clin Oncol 1996; **14**(9): 2527-39.
7. Bartelink H, Roelofsen F, Eschwege F, et al. Concomitant radiotherapy and chemotherapy is superior to radiotherapy alone in the treatment of locally advanced anal cancer: Results of a phase III randomized trial of the European organization for research and treatment of cancer radiotherapy and gastrointestinal cooperative groups. Journal of Clinical Oncology 1997; **15**(5): 2040-9.
8. Downing A, Morris EJA, Aravani A, et al. The effect of the UK coordinating centre for cancer research anal cancer trial (ACT1) on population-based treatment and survival for squamous cell cancer of the anus. Clinical Oncology 2015; **27**(12): 708-12.
9. Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FI, Trotti AI. AJCC Cancer Staging Manual New York: Springer; 2009.
10. James RD, Glynne-Jones R, Meadows HM, et al. Mitomycin or cisplatin chemoradiation with or without maintenance chemotherapy for treatment of squamous-cell carcinoma of the anus (ACT II): A randomised, phase 3, open-label, 2×2 factorial trial. The Lancet Oncology 2013; **14**(6): 516-24.
11. Engstrom PF, Arnoletti JP, Benson AB, 3rd, et al. NCCN clinical practice guidelines in oncology. Anal carcinoma. J Natl Compr Canc Netw 2010; **8**(1): 106-20.
12. Geh I, Moran B, Karandikar S, et al. Association of Coloproctology of Great Britain & Ireland Management of Colorectal Cancer Guidelines [in press]. 2017.
13. Glynne-Jones R, Nilsson PJ, Aschele C, et al. Anal cancer: ESMO-ESSO-ESTRO clinical practice guidelines for diagnosis, treatment and follow-up. Radiotherapy and Oncology 2014; **111**(3): 330-9.
14. Ajani JA, Winter KA, Gunderson LL, et al. Fluorouracil, mitomycin, and radiotherapy vs fluorouracil, cisplatin, and radiotherapy for carcinoma of the anal canal: a randomized controlled trial. JAMA 2008; **299**(16): 1914-21.
15. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. New England Journal of Medicine 1985; **312**(25): 1604-8.
16. Guyot P, Ades AE, Ouwens M, Welton N. Enhanced secondary analysis of survival data: Reconstructing the data from published Kaplan-Meier survival curves. BMC Medical Research Methodology 2012; **12**.
17. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med 2002; **21**(4): 589-624.
18. Harbord RM, Egger M, Sterne JAC. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Statistics in Medicine 2006; **25**(20): 3443-57.

19. Caldarella C, Annunziata S, Treglia G, Sadeghi R, Ayati N, Giovanella L. Diagnostic performance of positron emission tomography/computed tomography using fluorine-18 fluorodeoxyglucose in detecting locoregional nodal involvement in patients with anal canal cancer: a systematic review and meta-analysis. ScientificWorldJournal 2014; **2014**: 196068.
20. Jones M, Hruby G, Solomon M, Rutherford N, Martin J. The Role of FDG-PET in the Initial Staging and Response Assessment of Anal Cancer: A Systematic Review and Meta-analysis. Ann Surg Oncol 2015; **22**(11): 3574-81.
21. Bentzen AG, Guren MG, Wanderas EH, et al. Chemoradiotherapy of anal carcinoma: survival and recurrence in an unselected national cohort. International Journal of Radiation Oncology, Biology, Physics 2012; **83**(2): e173-80.
22. Leon O, Guren M, Hagberg O, et al. Anal carcinoma - Survival and recurrence in a large cohort of patients treated according to Nordic guidelines. Radiotherapy and Oncology 2014; **113**(3): 352-8.
23. Salmon RJ, Fenton J, Asselain B. Treatment of epidermoid anal canal cancer. American Journal of Surgery 1984; **147**(1): 43-8.
24. Svensson C, Goldman S, Friberg B. Radiation treatment of epidermoid cancer of the anus. International Journal of Radiation Oncology, Biology, Physics 1993; **27**(1): 67-73.
25. Nilsson PJ, Svensson C, Goldman S, Ljungqvist O, Glimelius B. Epidermoid anal cancer: a review of a population-based series of 308 consecutive patients treated according to prospective protocols. International Journal of Radiation Oncology, Biology, Physics 2005; **61**(1): 92-102.
26. Grabenbauer GG, Kessler H, Matzel KE, Sauer R, Hohenberger W, Schneider IH. Tumor site predicts outcome after radiochemotherapy in squamous-cell carcinoma of the anal region: long-term results of 101 patients. Diseases of the Colon & Rectum 2005; **48**(9): 1742-51.
27. Das P, Bhatia S, Eng C, et al. Predictors and Patterns of Recurrence After Definitive Chemoradiation for Anal Cancer. International Journal of Radiation Oncology Biology Physics 2007; **68**(3): 794-800.
28. Widder J, Kastenberger R, Fercher E, et al. Radiation dose associated with local control in advanced anal cancer: retrospective analysis of 129 patients. Radiotherapy & Oncology 2008; **87**(3): 367-75.
29. Fraunholz I, Rabeneck D, Gerstein J, et al. Concurrent chemoradiotherapy with 5-fluorouracil and mitomycin C for anal carcinoma: are there differences between HIV-positive and HIV-negative patients in the era of highly active antiretroviral therapy? Radiotherapy & Oncology 2011; **98**(1): 99-104.
30. Wolff HA, Raus I, Jung K, et al. High-grade acute organ toxicity as a positive prognostic factor in primary radiochemotherapy for anal carcinoma. International Journal of Radiation Oncology, Biology, Physics 2011; **79**(5): 1467-78.
31. DeFoe SG, Beriwal S, Jones H, et al. Concurrent Chemotherapy and Intensity-modulated Radiation Therapy for Anal Carcinoma - Clinical Outcomes in a Large National Cancer Institute-designated Integrated Cancer Centre Network. Clinical Oncology 2012; **24**(6): 424-31.
32. Lestrade L, De Bari B, Montbarbon X, Pommier P, Carrie C. Radiochemotherapy and brachytherapy could be the standard treatment for anal canal cancer in elderly patients? A retrospective single-centre analysis. Medical Oncology 2013; **30**(1): 402.
33. Chuong MD, Hoffe SE, Weber J, et al. Outcomes of anal cancer treated with definitive IMRT-based chemoradiation. Journal of Radiation Oncology 2012; **1**(2): 165-72.
34. Kim HJ, Huh JW, Kim CH, et al. Long-term outcomes of chemoradiation for anal cancer patients. Yonsei Medical Journal 2013; **54**(1): 108-15.
35. Eng C, Chang GJ, You YN, et al. Long-term results of weekly/daily cisplatin-based chemoradiation for locally advanced squamous cell carcinoma of the anal canal. Cancer 2013; **119**(21): 3769-75.
36. Fakhrian K, Sauer T, Klemm S, et al. Radiotherapy with or without chemotherapy in the treatment of anal cancer: 20-year experience from a single institute. Strahlentherapie und Onkologie 2013; **189**(1): 18-25.

37. Toh E, Wilson J, Sebag-Montefiore D, Botterill I. Neutrophil:lymphocyte ratio as a simple and novel biomarker for prediction of locoregional recurrence after chemoradiotherapy for squamous cell carcinoma of the anus. Colorectal Disease 2014; **16**(3): O90-7.
38. Yeung R, McConnell Y, Roxin G, et al. One compared with two cycles of mitomycin C in chemoradiotherapy for anal cancer: Analysis of outcomes and toxicity. Current Oncology 2014; **21**(3): 449-56.
39. De Bari B, Buglione M, Maddalo M, et al. External beam radiotherapy +/- chemotherapy in the treatment of anal canal cancer: a single-institute long-term experience on 100 patients. Cancer Investigation 2014; **32**(6): 248-55.
40. Oblak I, Cesnjevar M, Anzic M, et al. The impact of anaemia on treatment outcome in patients with squamous cell carcinoma of anal canal and anal margin. Radiology and Oncology 2016; **50**(1): 113-20.
41. Glynne-Jones R, Sebag-Montefiore D, Adams R, et al. Prognostic factors for recurrence and survival in anal cancer: generating hypotheses from the mature outcomes of the first United Kingdom Coordinating Committee on Cancer Research Anal Cancer Trial (ACT I). Cancer 2013; **119**(4): 748-55.
42. Benson Iii AB, Arnoletti JP, Bekaii-Saab T, et al. Anal carcinoma, version 2.2012: Featured updates to the NCCN guidelines. JNCCN Journal of the National Comprehensive Cancer Network 2012; **10**(4): 449-54.
43. Péley G, Farkas E, Sinkovics I, et al. Inguinal sentinel lymph node biopsy for staging anal cancer. Scandinavian Journal of Surgery 2002; **91**(4): 336-8.
44. Perera D, Pathma-Nathan N, Rabbitt P, Hewett P, Rieger N, Cummings BJ. Sentinel node biopsy for squamous-cell carcinoma of the anus and anal margin. Diseases of the Colon and Rectum 2003; **46**(8): 1027-31.
45. Chee KG, Nguyen DV, Brown M, Gandara DR, Wun T, Lara PN. Positron emission tomography and improved survival in patients with lung cancer: The Will Rogers phenomenon revisited. Archives of Internal Medicine 2008; **168**(14): 1541-9.
46. Seute T, Leffers P, Ten Velde GPM, Twijnstra A. Detection of brain metastases from small cell lung cancer: Consequences of changing imaging techniques (CT versus MRI). Cancer 2008; **112**(8): 1827-34.
47. Albertsen PC, Hanley JAH, Barrows GH, et al. Prostate cancer and the Will Rogers phenomenon. Journal of the National Cancer Institute 2005; **97**(17): 1248-53.
48. D'Souza G, Dempsey A. The role of HPV in head and neck cancer and review of the HPV vaccine. Prev Med 2011; **53 Suppl 1**: S5-S11.
49. Keane FK, Chen YH, Neville BA, et al. Changing prognostic significance of tumor stage and nodal stage in patients with squamous cell carcinoma of the oropharynx in the human papillomavirus era. Cancer 2015; **121**(15): 2594-602.
50. Egger M, Schneider M, Smith GD. Meta-analysis spurious precision? meta-analysis of observational studies. British Medical Journal 1998; **316**(7125): 140-4.
51. Renehan AG, Saunders MP, Schofield PF, O'Dwyer ST. Patterns of local disease failure and outcome after salvage surgery in patients with anal cancer. British Journal of Surgery 2005; **92**(5): 605-14.
52. Glynne-Jones R, Sebag-Montefiore D, Meadows HM, et al. Best time to assess complete clinical response after chemoradiotherapy in squamous cell carcinoma of the anus (ACT II): a post-hoc analysis of randomised controlled phase 3 trial. The Lancet Oncology 2017; **18**(3): 347-56.
53. PLATO_trial. PersonaLising Anal cancer radioTherapy dOse – Incorporating ACT3, ACT4 and ACT5 http://medhealth.leeds.ac.uk/info/430/solid_tumours/2210/plato [accessed 30 Dec 2016].