

# Is the 8th edition of the Union for International Cancer Control staging of oral cancer good enough?

D. Vuity<sup>a,\*</sup>, J. McMahon<sup>a</sup>, S. Takhiuddin<sup>a</sup>, C. Slinger<sup>a</sup>, D. McLellan<sup>b</sup>, C. Wales<sup>a</sup>, C. MacIver<sup>a</sup>, E. Thomson<sup>c</sup>, J. McCaul<sup>a</sup>, S. Hislop<sup>d</sup>, C. Lamb<sup>e</sup>, E. Stalker<sup>a</sup>, D. Young<sup>f</sup>

<sup>a</sup> Department of Oral and Maxillofacial Surgery, Queen Elizabeth University Hospital, 1345 Govan Road, Glasgow G51 4TF, UK <sup>b</sup> Department of Pathology, Queen Elizabeth University Hospital, 1345 Govan Road, Glasgow G51 4TF, UK <sup>c</sup> Department of Oral and Maxillofacial Surgery, Forth Valley Royal Hospital, Stirling Rd, Larbert FK5 4WR, UK <sup>d</sup> Department of Oral and Maxillofacial Surgery, University Hospital Crosshouse, Kilmarnock Rd, Crosshouse, Kilmarnock KA2 0BE, UK <sup>e</sup> Department of Radiotherapy, Beatson West of Scotland Cancer Centre, 1053 Great Western Rd, Glasgow G12 0YN, UK <sup>f</sup> Department of Mathematics and Statistics, University of Strathclyde, Livingstone Tower, 26 Richmond St, Glasgow G1 1XH, UK

## Abstract

Depth of invasion is an important predictor of survival. A study by the International Consortium (ICOR) for Outcome Research proposed incorporation of it (together with the greatest surface dimension, or the anatomical criteria, or both) into the T stage. This has been adopted in part by the 8th edition of the Union for International Cancer Control (UICC) TNM 8 classification of malignant tumours for oral squamous cell carcinoma (SCC). Our aim was to verify depth of invasion as an independent prognostic factor, and to validate the staging by comparing it with that specified in the 7th edition (TNM 7) and the T-staging model proposed by the International Consortium. We retrospectively studied 449 patients who had had operations for a previously untreated primary oral cancer between 2006 and 2014 at a single centre, and analysed the independent predictive value of depth of invasion for both disease-specific and overall survival. It was an independent predictor of disease-specific survival as were sex, perineural invasion, and N stage. It was also an independent predictor of overall survival together with sex and N status. Staging in TNM 8 gave a better balance of distribution than that in TNM 7, but did not discriminate between prognosis in patients with T3 and T4 disease. The proposed International Consortium rules for T-staging gave an improved balance in distribution and hazard discrimination. The incorporation of depth of invasion into the T-staging rules for oral SCC improved prognostic accuracy and is likely to influence the selection of patients for adjuvant treatment. Our findings suggest that the TNM 8 staging lacks hazard discrimination in patients with locally-advanced disease because its T4 staging is restricted to anatomical criteria.

## Introduction

TNM staging for cancer was first developed and published by Pierre Denoix in 1952,<sup>1</sup> and the first edition of the Union Internationale Contre le Cancer (UICC) TNM Classification of Malignant Tumours was published in 1953.<sup>2</sup> The American Joint Committee on Cancer (AJCC), which was established in 1959, published the first cancer staging manual in 1977.<sup>3</sup> Close cooperation between these organisations enabled the development of a universal system for the classification of tumours of epithelial origin.

Tumour thickness or depth of invasion, or both, have been reported to be objective predictors of nodal involvement and survival,<sup>4</sup> and proposed modifications of the T staging of oral squamous cell carcinoma (SCC) incorporate depth of invasion.<sup>5,6</sup>

In 2014, the International Consortium (ICOR) for Outcome Research in head and neck cancer investigated the predictive value of depth of invasion.<sup>7</sup> This retrospective multicentre study of 3149 patients from 11

“comprehensive cancer centres” worldwide showed that depth of invasion was predictive of disease-specific survival. Based on this, a new model was proposed (Table 1) that incorporated depth of invasion into the T-staging criteria. It was also incorporated into the 8th edition of the UICC T staging of oral SCC (TNM 8),<sup>8</sup> which is similar to that proposed by the ICOR, but T4 staging is restricted to anatomical criteria.

Our aim was to verify depth of invasion as an independent prognostic factor, and to validate staging in the 8th edition (TNM 8) by comparing it with that specified in the 7th edition (TNM 7) and the T-staging model proposed by the ICOR.

## Methods

We retrospectively studied 467 previously untreated patients with oral SCC, who were treated with curative intent by operation, with or without adjuvant treatment, between 2006 and 2014 (inclusive) at our hospital. We excluded 18 (4%) because of insufficient information on depth of invasion, so the final group comprised 449 patients. Primary tumours were excised with a macroscopic margin of 1 cm based on clinical examination and cross-sectional imaging, and patients were followed up for a minimum of 24 months. Clinical and pathological data points were retrieved from case records.

Tables 2A and 2B show the clinicopathological variables. We used the definitive histopathology report on the resected specimen to allocate pathological T stages according to TNM 7, TNM 8, and ICOR criteria. Tumour less than 1 mm from a resection margin was considered involved; tumour 1–5 mm from the closest margin, close.

### Histopathology

Moore et al<sup>9</sup> described the depth of invasion as the distance from the epithelial basal membrane to the deepest point of infiltration. We reconstructed the epithelial surface if it was ulcerated or the tumour was exophytic. Pathologists who had a subspecialty interest in diseases of the head and neck did the histological evaluation.

### Statistical analysis

Statistical analysis was done with the help of SPSS Statistics for Windows version 22.0 (IBM Corp). The clinical end points of interest were overall survival and disease-specific survival. Survival time was calculated from the date of operation. A significance level of 5% was used for all statistical analyses. For continuous variables (depth of invasion, age) univariate analysis was done by comparison of means using the *t* test or Mann-Whitney U test as appropriate (Fig. 1B). For categorical variables, including T stage, Kaplan-Meier survival curves were constructed and the significance calculated with the log-rank test (Cox-Mantel). All clinicopathological factors that were predictive of survival (including depth of invasion) on univariate analysis were entered into a Cox regression model to establish independent predictors of survival (Table 3). Cox's proportional hazard regression was used to calculate the separation between the survival curves in each system with the T4 category as the comparator. The T staging systems were also evaluated using the  $-2 \log$  likelihood test for fit of the data.

## Results

There were 180 women and 269 men (ratio 1:1.5) with a median (range) age of 62 (22–94) years. Median (range) follow up was 51(2–60) months. The distribution of patients in each T category according to TNM 7, TNM 8, and the proposed ICOR criteria, is shown in Table 2A. TNM 7 rules resulted in a comparatively poor distribution balance across the T categories with a paucity of patients in T3 (n = 14).

Specimens showed perineural invasion (n = 107, 24%), lymphovascular invasion (n = 105, 23%), and extranodal extension (n = 93, 21%). A total of 33 patients (7%) had involved margins. N stages are shown in Table 2A. These were pathological (pN) if patients had had lymphadenectomy when the primary tumour was resected.

The mean (range) depth of invasion was 8.16 (0.3-40)mm. On univariate analysis (Table 2B) it was a predictor ( $p < 0.0001$ ) of both disease-specific and overall survival. On multivariate analysis it was an independent predictor of disease-specific survival, as were sex, perineural invasion, and regional lymph node metastases. It was an independent predictor of overall survival together with sex and pathological nodal involvement (Table 3). Disease specific survival curves for the three staging systems are shown in Fig. 1 and the overall survival curves in Fig. 2.

Both show comparatively poor hazard discrimination between T3 and T4 groups for both TNM 7 and 8 rules. Cox's regression analysis (Table 4) showed that only the ICOR T-staging rules could discriminate between the prognosis for patients with T3 primary tumours and those with T4 disease. Goodness-of-fit tests showed that TNM 8 improved performance more than TNM 7, but was inferior to the ICOR staging system (Table 4).

## Discussion

Accurate staging is key to the management of oral SCC, as it informs treatment and provides important information that can help clinicians, patients, and their families understand the probability of a cure. The status of the regional lymph nodes is the most important determinant of outcome and strongly influences the recommendation for adjuvant treatment after primary operation, but T staging has not consistently been independently prognostic.<sup>10</sup> It is possible that its relatively low discriminatory value is, at least partly, a consequence of the staging rules applied.

This has been recognised in the latest (8th) edition of the UICC TNM staging manual. Two key changes have been made. One is the omission of an anatomical criterion (which has caused much vexed discussion in multidisciplinary team meetings), specifically the allocation of T4 status based on involvement of deep extrinsic muscles of the tongue. However, it is not clear what this means or whether it indicates advanced disease or an accident of anatomy. A tumour in the midline floor of the mouth has to invade only a few millimetres before it penetrates the oral surface of genioglossus. Similarly, a tumour on the posterolateral tongue may invade the styloglossus, palatoglossus, or hyoglossus, as soon as it penetrates much beyond the submucosa. These accidents of anatomy do not inform our understanding of tumour biology, and this problem has almost certainly resulted in too many patients being staged as T4, which has devalued its use for staging of the primary site. TNM 8 has continued to use only anatomical rules for the allocation of T4, which might explain its inferior performance in hazard discrimination compared with the proposed model by the ICOR (Figs. 1 and 2).

Involvement of the skin might also be an accident of anatomy. Any tumour that is deeper than 5–6 mm and arises on the buccal mucosa close to the oral commissure is likely to involve skin, but we have been unable to find evidence that this in itself carries an adverse prognosis. Similarly, oral SCC that arises on the gum will invade the mandible relatively early. We speculate that the stipulation in the ICOR model that T4 status requires a depth of invasion of 10 mm or more, overcomes this potential weakness.

Our findings that depth of invasion is an independent predictor of both disease-specific and overall survival further validate the recognition of its importance in the 8th edition. It has resulted in a more balanced allocation to the T categories and greater prognostic accuracy. Recent trial data show that clinical teams have already realised that a 10 mm depth of invasion is an indication for adjuvant radiotherapy.<sup>11</sup> The allocation of T3 status to patients with oral SCC that was previously staged as T1 or T2, will broaden that practice and improve the balance in the distribution between T categories. There are, however, difficulties. Histopathological assessment of depth of invasion is subjective and the interobserver error is unknown. Depth of invasion is also a continuous variable

that will probably be associated with a continuum of risk (albeit non-linear), so the dichotomous categorisation at 5mm and 10mm<sup>12</sup> will probably be challenged. Nevertheless, our findings support those of the much larger study by the ICOR.<sup>7</sup>

Optimal staging rules should meet three criteria: hazard consistency (same survival rate within each stage when applied to different groups of patients), hazard discrimination (different survival rates among groups), and balance in distribution (enough patients in each group to predict survival (statistical strength)).<sup>13</sup> The survival curves for the different T staging systems (Figs. 1 and 2) show that TNM 8 offers a better balance in distribution than TNM 7. However, the TNM 8 rules in this group did not separate the survival curves for the T3 and T4 categories, possibly for the reasons outlined above, and they therefore lack hazard discrimination. This contradicts the findings in the two-centre validation group<sup>13</sup> that suggested that TNM 8 might lack hazard consistency. In contrast, we found that the proposed ICOR staging system showed good balance in distribution and good hazard discrimination.

In conclusion, application of the staging rules in TNM 8 and those of the ICOR resulted in a better balance in the distribution of patients across the T stages than application of those in TNM 7. However, only the ICOR rules gave adequate hazard discrimination in this group, which suggests that accurate staging of T4 requires consideration of depth of invasion as well as anatomical features. This finding now requires confirmation in larger retrospective and prospective studies.

## References

1. Denoix PF. Nomenclature and classification of cancers based on an atlas. *Acta Unio Int Contra Cancrum* 1953;**9**:769–71 (Undetermined language).
2. Committee of TNM classification of the UICC. *Acta Oncol (Madr)* 1968;**7**:254–64 (paper in Spanish).
3. American Joint Committee for Cancer Staging and End Results Reporting. *Manual for staging of cancer 1977*. Chicago: American Joint Committee; 1977.
4. Pentenero M, Gandolfo S, Carrozzo M. Importance of tumor thickness and depth of invasion in nodal involvement and prognosis of oral squamous cell carcinoma: a review of the literature. *Head Neck* 2005;**27**:1080–91.
5. Yuen AP, Lam KY, Wei I, et al. A comparison of the prognostic significance of tumor diameter, length, width, thickness, area, volume, and clinicopathological features of oral tongue carcinoma. *Am J Surg* 2000;**180**:139–43.
6. Howaldt HP, Kainz M, Euler B, et al. Proposal for modification of the TNM staging classification for cancer of the oral cavity. DOSAK. *J Craniomaxillofac Surg* 1999;**27**:275–88.
7. International Consortium for Outcome Research (ICOR) in Head and Neck Cancer, Ebrahimi A, Gil Z, et al. Primary tumor staging for oral cancer and a proposed modification incorporating depth of invasion: an international multicenter retrospective study. *JAMA Otolaryngol Head Neck Surg* 2014;**140**:1138–48.
8. Brierley JD, Gospodarowicz MK, Wittekind C, editors. *TNM classification of malignant tumours*. eighth edition Chichester: John Wiley & Sons; 2017.
9. Moore C, Kuhns JG, Greenberg RA. Thickness as prognostic aid in upper aerodigestive tract cancer. *Arch Surg* 1986;**121**:1410–4.
10. Rogers SN, Brown JS, Woolgar JA, et al. Survival following primary surgery for oral cancer. *Oral Oncol* 2009;**45**:201–11.
11. D’Cruz AK, Vaish R, Kapre N, et al. Elective versus therapeutic neck dissection in node-negative oral cancer. *N Engl J Med* 2015;**373**:521–9.

12. Lydiatt WM, Patel SG, O'Sullivan B, et al. Head and neck cancers- major changes in the American Joint Committee on cancer eighth edition cancer staging manual. *CA Cancer J Clin* 2017;**67**:122–37.
13. Groome PA, Schulze KM, Mackillop WJ, et al. A comparison of published head and neck stage groupings in carcinomas of the tonsillar region. *Cancer* 2001;**92**:1484–94.

Table 1

Criteria of the different staging systems.

Staging System

- A. International Consortium:
  - T1 UICC 7th T1 or less in greatest dimension, maximum depth of invasion <5 mm
  - T2 UICC 7th T1, maximum depth of invasion  $\geq$ 5 mm
  - T3 UICC 7th T2, maximum depth of invasion  $\geq$ 10mm
  - T4 UICC 7th T3-T4, maximum depth of invasion  $\geq$ 10mm
  
- B. 8th edition UICC (TNM8)
  - T1 2 cm or less in greatest dimension, maximum depth of invasion DOI <5 mm
  - T2 2 cm or less in greatest dimension, maximum depth of invasion  $\geq$ 5 mm, or tumour >2 cm and 4 cm or more in greatest dimension, maximum depth of invasion <10 mm
  - T3 Any tumour maximum depth of invasion  $\geq$ 10 mm or tumour >4 cm in greatest dimension
  - T4a Invades through cortical bone, or maxillary sinus, or invades skin
  - T4b Invades masticator space, pterygoid plates or skull base, or encases the internal carotid artery
  
- C. 7th edition UICC (TNM7)
  - T1 2 cm or less in greatest dimension
  - T2 Tumour >2 cm and 4 cm or more in greatest dimension
  - T3 Tumour >4 cm in greatest dimension
  - T4a Invades bone, or maxillary sinus, or invades skin, into deep extrinsic muscles of the tongue
  - T4b Invades masticator space, pterygoid plates or skull base, or encases the internal carotid artery

Table 2A

Descriptive and univariate analysis of the categorical data.

	No. (%)	Disease-specific survival	Overall survival
Clinicopathological variables: Sex:		0.015	0.009
Female	180 (40)		
Male	269 (60)		
pT category 7th edition UICC (TNM 7):		<0.0001	<0.0001
T1	200 (45)		
T2	111 (25)		
T3	14 (3)		
T4	124 (27)		
pT category 8th edition UICC (TNM 8):		<0.0001	<0.0001
T1	143 (32)		
T2	136 (30)		
T3	99 (22)		
T4	71 (16)		
pT category International Consortium:		<0.0001	<0.0001
T1	142 (32)		
T2	130 (29)		
T3	82 (18)		
T4	95 (21)		
Pathological N category:		<0.0001	<0.0001
N0	282 (63)		
N1	53 (12)		
N2	113 (25)		
N3	1 (0.2)		
Lymphovascular invasion:		<0.0001	<0.0001
Present	105 (23)		

Absent	344 (77)		
Perineural invasion:		<0.0001	<0.0001
Present	107 (24)		
Absent	342 (76)		
Extranodal extension:		<0.0001	<0.0001
Present	93 (79)		
Absent	356 (21)		
Margins:		0.825	0.348
Involved	33 (7)		
Close	233 (52)		
Clear	183 (41)		
Adjuvant treatment:			
Radiotherapy	97 (22)		
Neoadjuvant chemotherapy	15 (3)		
Concurrent chemoradiotherapy	86 (19)		

---



Table 2B  
 Continuous clinicopathological variables

	Mean (SD)	No.	p value
Age (years):			
Alive	60.72(11.7)	309	0.082
Died any cause	62.76 (10.9)	140	
Alive	61.37 (11.6)	364	0.947
Died of disease	61.28 (11.06)	85	
Depth of invasion (mm):			
Alive	6.52(5.12)	309	<0.0001
Died any cause	11.78 (8.37)	140	
Alive	6.9 (5.7)	364	<0.0001
Died of disease	13.54 (8.32)	85	

Table 3

Independent predictors of survival.

Variable	B (SE)	Wald Statistic	p value	Exp(B) (95% CI)
Overall survival:				
Sex	0.383 (0.186)	4.213	0.040	1.47 (1.017 to 2.113)
Depth of invasion	0.068 (0.010)	43.541	<0.0001	1.07 (1.049 to 1.092)
N stage	0.478 (0.095)	25.076	<0.0001	1.61 (1.338 to 1.945)
Disease-specific survival:				
Sex	0.519 (0.246)	4.441	0.035	1.681 (1.037 to 2.724)
Depth of invasion	0.073 (0.013)	32.766	<0.0001	1.075 (1.049 to 1.102)
Perineural invasion	0.556 (0.231)	5.792	0.016	1.744 (1.109 to 2.744)
N stage	0.709 (0.132)	29.062	<0.0001	2.032 (1.570 to 2.629)

Table 4

Cox proportional hazard regression of the Kaplan Meier curves in each staging system. (baseline T4), and -2 log likelihood goodness-of-fit testing of the different staging systems.

	TNM 7	TNM 8	ICOR
Cox's regression:			
Disease-specific survival:			
T4 - T1	p < 0.0001	p < 0.0001	p < 0.0001
T4 - T2	p = 0.009	p < 0.0001	p < 0.0001
T4 - T3	p = 0.332	p = 0.701	p = 0.009
Overall survival:			
T4 - T1	p < 0.0001	p < 0.0001	p < 0.0001
T4 - T2	p = 0.002	p < 0.0001	p < 0.0001
T4 - T3	p = 0.390	p = 0.758	p = 0.001
Goodness of fit of the different staging systems:			
Disease-specific survival	964.007	951.683	946.452
Overall survival	1605.68	1593.729	1584.867

ICOR: International Consortium for Outcome Research.

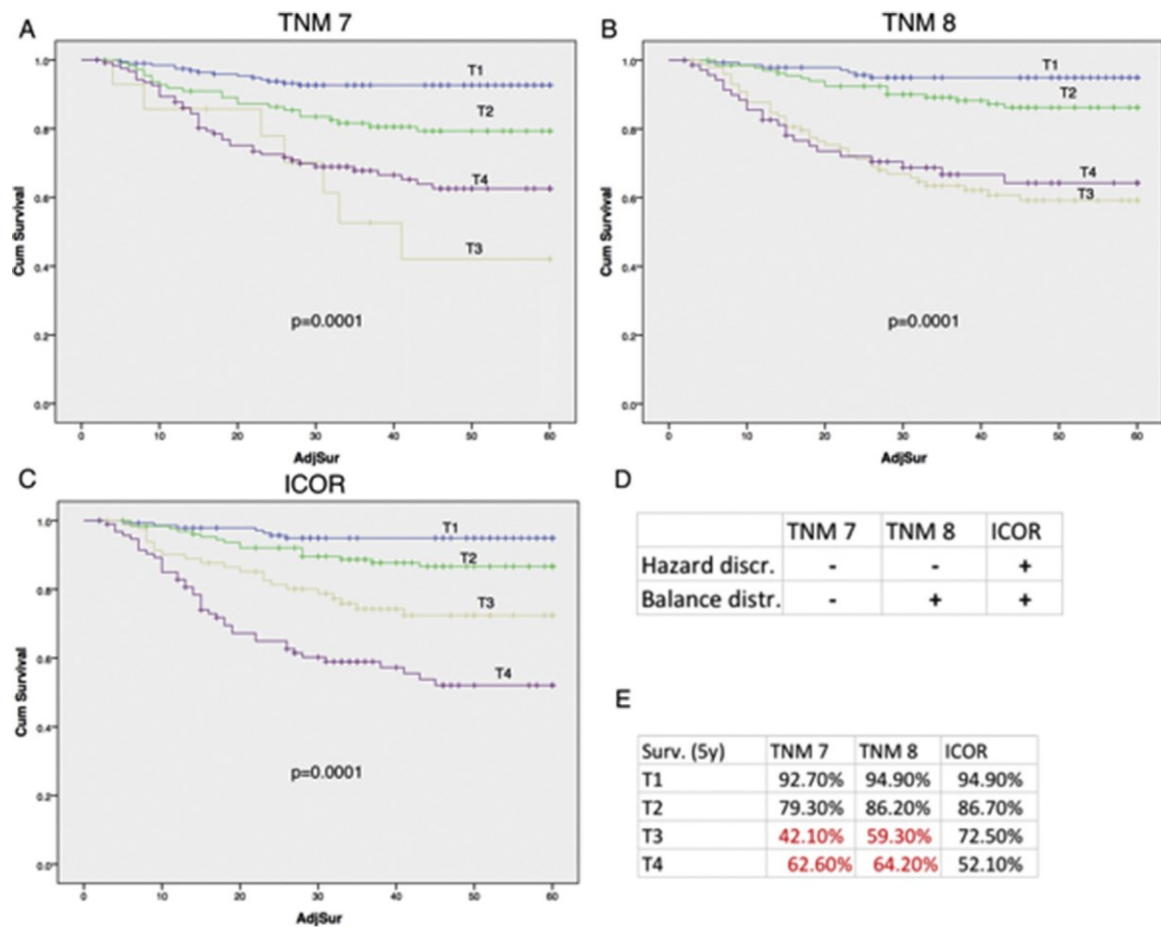


Figure 1. Disease-specific survival calculated with Kaplan Meier curves (A = TNM 7; B = TNM 8; C = International Consortium staging; D = presence of the examined criteria in each staging system: hazard discrimination and balance in distribution; E = five-year survival according to TNM 7, TNM 8, and ICOR staging).

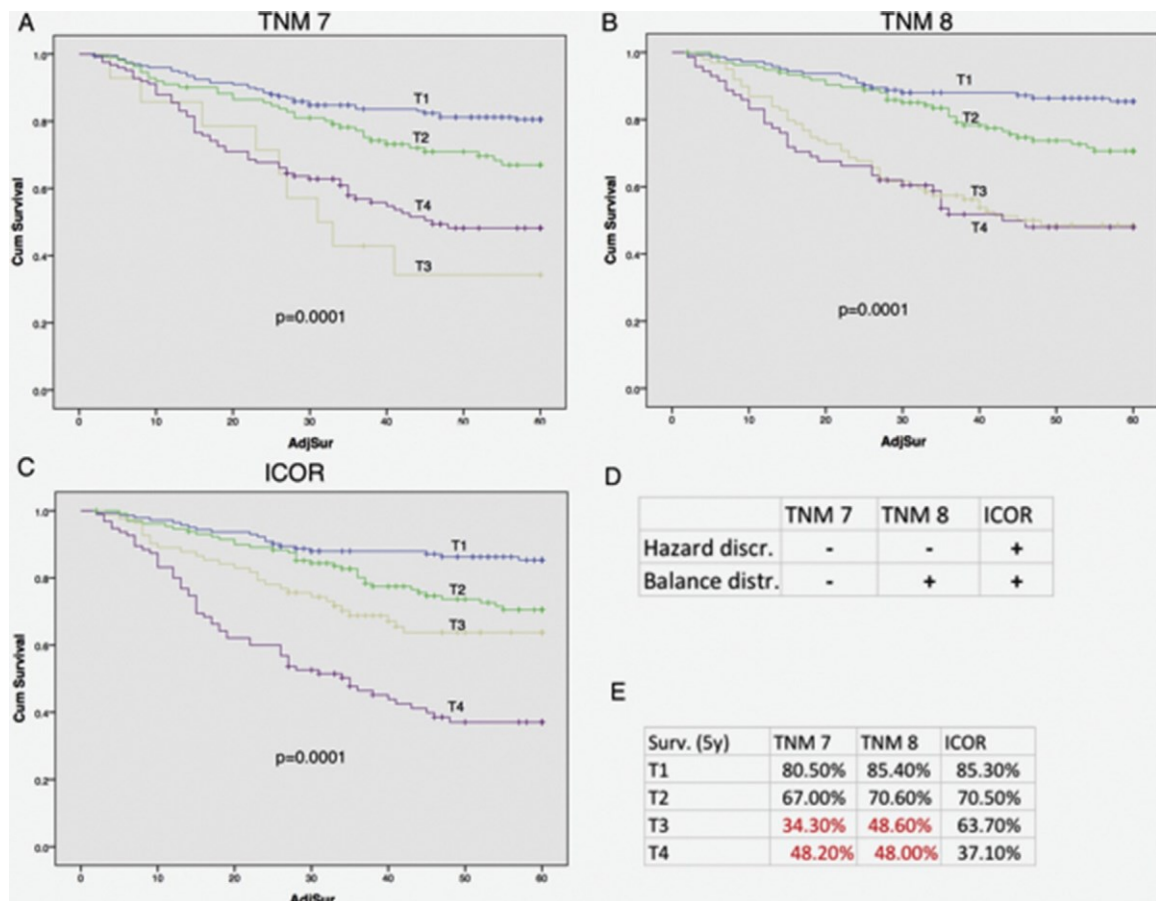


Figure 2. Overall survival calculated with Kaplan Meier curves (A = TNM 7; B = TNM 8; C = International Consortium staging; D = presence of the examined criteria in each staging system: hazard discrimination and balance in distribution; E = five-year survival according to TNM 7, TNM 8, and ICOR staging).