Université de Montréal

L'évaluation préopératoire de la profondeur d'invasion des carcinomes épidermoïdes de la langue mobile

Connaissances actuelles et rôle diagnostique de la biopsie au poinçon

Par

Béatrice Voizard, MD, FRCSC

Faculté de Médecine

Mémoire présenté en vue de l'obtention du grade de Maîtrise en Sciences Biomédicales, option médecine expérimentale

Août 2022

Université de Montréal

Faculté de Médecine

Ce mémoire intitulé

L'évaluation préopératoire de la profondeur d'invasion des carcinomes épidermoïdes de la langue mobile

Connaissances actuelles et rôle diagnostique de la biopsie au poinçon

Présenté par

Béatrice Voizard, MD, FRCSC

A été évalué par un jury composé des personnes suivantes

Olivier Abboud, MD, MSc, FRCSC
Président-rapporteur

Apostolos Christopoulos, MD, MSc, FRCSC
Directeur de recherche

Philippe Lavigne, MD, FRCSC
Membre du jury

Résumé

L'inclusion récente de la profondeur d'invasion (PI) dans la classification des carcinomes épidermoïdes de la cavité orale de l'American Joint Committee on Cancer (AJCC) a des répercussions cliniques majeures. Plusieurs études ont récemment évalué la fiabilité de diverses modalités d'imagerie et techniques de biopsie pour mesurer la PI en préopératoire. L'objectif premier de ce mémoire est de réviser systématiquement la littérature et comparer les différentes méthodes décrites de mesure de PI en préopératoire pour les carcinomes épidermoïdes de la langue mobile. Le second objectif est d'étudier la précision et la fiabilité de la mesure de PI sur une biopsie au poinçon dans les carcinomes épidermoïdes de stade in situ (Tis)-T1-T2, N0 de la langue mobile.

Une revue systématique a été effectuée en suivant le guide PRISMA[1]. Les études évaluant la fiabilité de la PI mesurée sur la biopsie ou l'imagerie médicale, en les comparant à la PI histopathologique finale, ont été inclues dans une méta-analyse afin d'obtenir des coefficients de corrélation combinés pour chaque modalité d'imagerie. L'imagerie par résonnance magnétique (IRM) s'est avérée être la modalité d'imagerie la mieux étudiée et présente une bonne fiabilité. Le *computed tomography* (CT) scan est peu étudié, mais semble moins fiable. L'échographie linguale ne peut être comparée à ces deux modalités d'imagerie car elle est plus fréquemment utilisée pour mesurer l'épaisseur tumorale que la PI.

La seconde étude est une preuve de concept prospective. Un poinçon profond a été utilisé pour échantillonner la portion la plus profonde de carcinomes épidermoïdes de la langue mobile de stade Tis-T1-2, N0 chez 27 patients. Des coefficients de corrélation de Spearman ont été calculés entre la PI estimée à la palpation manuelle, mesurée à la biopsie, et à l'histopathologie. La sensibilité et la spécificité de la biopsie au poinçon pour distinguer le Tis du carcinome épidermoïde invasif ont été calculées. Bien que la PI mesurée à la biopsie ne corrèle pas fortement avec la PI histopathologique, cette preuve de concept est limitée par la taille d'échantillon. La biopsie au poinçon semble toutefois être un outil fiable pour distinguer le Tis de l'invasif. D'autres

études sont nécessaires avant de pouvoir recommander l'utilisation systématique de la biopsie pour décider en préopératoire si un évidement cervical électif est nécessaire.

Mots-clés : Cancer cervico-facial, Cancer de la cavité orale, Cancer de la langue, Carcinome épidermoïde, Biopsie, Imagerie médicale, Gradation tumorale, Invasion Tumorale

Abstract

The inclusion of depth of invasion (DOI) in the American Joint Committee on Cancer's staging system for oral cavity squamous cell carcinoma (SCC) has major clinical impacts. Recent studies have evaluated the reliability of imaging modalities and biopsy techniques to measure DOI preoperatively. The first objective of this master's thesis is to systematically review and compare the preoperative DOI measurement methods that have been studied so far in oral tongue SCC (OTSCC). The second objective is to prospectively study the precision and reliability of punch biopsy to measure DOI preoperatively in early (in situ (Tis)-T1-T2, N0) OTSCC, and its ability to distinguish Tis from invasive carcinoma.

A systematic review was conducted according to the PRISMA guidelines. Studies that evaluated the reliability of DOI measured on biopsy or imaging (rDOI) by comparing it to DOI on histopathology (pDOI) were included in a meta-analysis to obtain pooled correlation coefficients for each imaging modality. Overall, magnetic resonance imaging (MRI) is the better studied modality. It has a good reliability to measure preoperative rDOI in OTSCC. CT is less studied but appears to be less reliable. Ultrasound (US) cannot be compared to these imaging modality as it has been used more often to measure tumor thickness (TT) than DOI.

The second study is a prospective proof-of-concept. A deep punch biopsy was used to sample tumors preoperatively in the deepest part of the tumor in 27 patients with early (Tis-T1-2, N0) oral tongue squamous cell carcinoma. Spearman's correlations were calculated between DOI measured on digital palpation (cDOI), biopsy (bDOI) and final pDOI. The sensitivity and specificity of punch biopsy to distinguish Tis from invasive carcinoma was also calculated. Although bDOI does not seem to correlate strongly with pDOI, this proof-of-concept was limited by a small sample size. Punch biopsy appears to be a reliable tool to distinguish Tis from invasive carcinoma.

Further studies on punch biopsy are needed to recommend its use to evaluate pDOI preoperatively and determine whether elective neck dissection is necessary in early OTSCC.

Keywords : Head and Neck Cancer, Oral cancer, Tongue Neoplasms, Squamous Cell Carcinoma, Biopsy, Diagnostic Imaging, Neoplasm Staging, Neoplasm Invasiveness

Table des matières

Liste des tableaux

Liste des figures

Liste des sigles et abréviations

AJCC : American Joint Committee on Cancer

AJCC7: 7ème édition de la classification TNM du AJCC

AJCC8: 8ème édition de la classification TNM du AJCC/ eighth edition of the American Joint Committee on Cancer

bDOI : depth of invasion measured on biopsy

BM: basement membrane

cDOI : depth of invasion measured on digital palpation / clinical depth of invasion

CHUM: Centre Hospitalier de l'Université de Montréal

$CI_{95\%}$: 95% confidence interval

CT: computed tomography

DOI: depth of invasion

END: elective node dissection

ICC: intraclass coefficient

IRM : imagerie par résonance magnétique

LVI : lymphovascular invasion

MRI: magnetic resonance imaging

NCCN : National Comprehensive Cancer Network

OR : operating room

OTSCC: oral tongue squamous cell carcinoma

pDOI: depth of invasion measured on histopathology

PNI : perineural invasion

PPV : positive predictive value

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses

pTT: tumor thickness measured on histopathology

rDOI: depth of invasion measured on imaging

rTT: tumor thickness measured on imaging

SCC: squamous cell carcinoma

SRMA: systematic review and meta-analysis

Tis : Carcinome in situ / Carcinoma in situ

TNM: Tumor, Node, Metastatis

TT: tumor thickness

UICC: Union Internationale Contre le Cancer

US: ultrasound

À Alain, dont j'admirerai toujours la curiosité insatiable

À Marie-Paule, mon modèle de ténacité infinie

À Julien, ma plus grande certitude

Remerciements

Merci à Dr Apostolos Christopoulos, qui m'a suivie durant cette maîtrise allongée par une pandémie et une résidence chirurgicale prenante. Merci de m'avoir crue lorsque j'ai promis que je terminerais tout ce que j'avais commencé.

Merci également à tous les co-auteurs des deux articles.

***

Merci à tous les professeurs qui ont mis en lumière les profondeurs de mon ignorance et ont, par le fait-même, cultivé ma curiosité.

***

Merci à ma famille et mes amis. Je chercherai toujours les limites maintes fois repoussées de mon imaginaire, mais même au-delà de celles-ci, je ne saurai jamais m'imaginer mieux entourée.

Chapitre 1 – Introduction

Les néoplasies oto-rhino-laryngologiques sont fréquentes, représentant 5.7 % de la mortalité attribuable au cancer globalement.[2] Parmi celles-ci, environ 25% se développement au niveau de la cavité orale. La langue mobile est le sous-site de la cavité orale le plus fréquemment touché par ces néoplasies.[3] On recensait au moins 1000 nouveaux cas de néoplasies de la langue mobile diagnostiqués au Canada en 2015 [4] alors qu'aux États-Unis, le nombre attendu de nouveaux diagnostics de cancers de la langue mobile était de plus de 17 000 en 2022.[3] L'établissement d'un diagnostic adéquat et la pronostication précise des néoplasies linguales représentent donc des enjeux majeurs dans la prise en charge chirurgicale et oncologique optimale des néoplasies de la sphère oto-rhino-laryngologique.

Récemment, le système de classification des tumeurs de l'American Joint Committee on Cancer (AJCC), a été révisé dans le but d'inclure de nouveaux facteurs pathologiques à la classification des tumeurs linguales.[5] L'importance pronostique de ces facteurs pathologiques a été démontrée dans la littérature scientifique contemporaine.

Ce mémoire de recherche brosse en premier lieu un portrait clair des défis cliniques découlant de l'inclusion de nouvelles mesures pathologiques dans la classification des cancers de la langue. En second lieu, une revue systématique de la littérature avec méta-analyse compare les différentes modalités diagnostiques étudiées jusqu'à présent pour classifier adéquatement les néoplasies linguales selon leur profondeur d'invasion, avant l'obtention de la mesure histopathologique finale de la tumeur. Finalement, l'utilité d'une méthode biopsique spécifique pour la pronostication adéquate préopératoire des néoplasies linguales est étudiée dans un deuxième article.

Chapitre 2 – Revue de littérature et problématique de recherche

2.1 Anatomie oncologique de la cavité orale

La cavité orale est un site de la tête et du cou dont les limites anatomiques selon l'AJCC sont en antérieur la jonction entre la peau et le vermillon labial et en postérieur la jonction entre les palais dur et mou, les papilles linguales circumvallées et les piliers amygdaliens antérieurs.[5] La cavité orale est divisée en huit sous-sites permettant de répertorier adéquatement l'origine anatomique et les sites d'envahissement des tumeurs. Ceux-ci sont les lèvres muqueuses, la muqueuse buccale, les crêtes alvéolaires inférieures et supérieures, le palais dur, les trigones rétro-molaires, le plancher buccal et la langue mobile. La langue mobile s'étend de son bout en antérieur jusqu'aux papilles circumvallées en postérieur. Elle est subdivisées en dorsum, bords latéraux et face ventrale.[5]

2.1.1 Anatomie microscopique de la langue mobile

La langue mobile est composée d'un épithélium de surface, d'une membrane basale et d'un ensemble sous-jacent de quatre muscles intrinsèques : deux muscles longitudinaux, supérieur et inférieur, un muscle vertical et un muscle horizontal. L'épithélium lingual de surface est kératinisant sur le dorsum, mais ne l'est pas sur les bords latéraux ainsi que sur la portion ventrale de la langue.

2.2 Le carcinome à cellules squameuses de la langue mobile

2.2.1 Physiopathologie et corrélations pathologiques

La physiopathologie du carcinome à cellules squameuses de la langue mobile débute par une altération des mécanismes de signalisation moléculaire intracellulaire entraînant l'accumulation de mutations cellulaires nommées atypies. Les atypies affectent soit l'anatomie microscopique intracellulaire ou encore l'architecture tissulaire de l'épithélium. Par exemple, une augmentation du ratio nucléocytoplasmique, une altération de l'aspect de la chromatine, une dédifférenciation

cellulaire ou encore une dyskératose de l'épithélium peuvent apparaître. Ces altérations microscopiques cellulaires ou architecturales constituent la dysplasie épithéliale. Celle-ci progresse de la couche basale de l'épithélium lingual vers sa surface superficielle. La sévérité de la dysplasie est classée comme légère, modérée ou sévère selon qu'elle touche l'épithélium dans son tiers profond, moyen ou sur toute son épaisseur.

La membrane basale épithéliale constitue une barrière physiologique à l'envahissement tumoral. Lorsque les atypies cellulaires et architecturales épithéliales sont assez sévères, mais qu'elles n'envahissent pas la membrane basale, la lésion est classifiée pathologiquement comme du carcinome épidermoïde in situ (Tis). Toutefois, lorsque les cellules dysplasiques envahissent la membrane basale vers le stroma musculaire sous-jacent, le carcinome épidermoïde est alors considéré comme envahissant. La prise en charge oncologique de ces deux pathologies, soit le Tis et l'envahissant, diffère grandement en raison du potentiel métastatique inexistant pour le premier et bien démontré du second.

2.2.2 Classification des tumeurs de la cavité orale selon la 8e édition du AJCC

La classification des cancers de la langue mobile est présentement en évolution ; la 8^{e} édition de la classification Tumor, Node, Metastasis (TNM) du AJCC (AJCC8), publiée en 2017 et implanté en 2018, considère la profondeur d'invasion (PI)[*] des cancers de la langue mobile comme un facteur déterminant dans ce *staging* servant à établir le pronostic et guider l'approche thérapeutique des néoplasies. [5] (Voir *Table 1*, p74)

2.2.3 Profondeur d'invasion tumorale

La PI est définie comme la profondeur tumorale allant de la membrane basale épithéliale à la marge profonde de la tumeur. Celle-ci est définie comme la limite entre le tissu néoplasique et le stroma musculaire sain sous-jacent. L'addition de cette variable dans AJCC8 permet de distinguer les tumeurs de large superficie, mais peu invasives, des tumeurs plus invasives de superficie moindre, qui peuvent présenter des comportements oncologiques distincts. Les

[*] La profondeur d'invasion se traduit en anglais comme *depth of invasion (DOI).*

tumeurs de PI < 5mm sont de stade T1, alors que celles d'une PI de 5mm à 10mm sont de stade T2 [6]. Quant à elles, les tumeurs de stade Tis n'impliquent que l'épithélium et ont donc une PI de valeur nulle.

Le potentiel de métastases est nul pour le Tis et des études ont même confirmé que le potentiel de métastases occultes à long terme est négligeable pour les tumeurs invasives de tous les sous-sites de la cavité orale de PI inférieure à 1 mm. [7]

2.2.4 Distinguer épaisseur tumorale et PI

Alors que l'épaisseur tumorale représente la taille totale de la tumeur et comprend à la fois la portion infiltrante profonde à la membrane basale ainsi que la portion exophytique au-dessus de la membrane basale, la profondeur d'invasion ne concerne que la portion infiltrante. L'épaisseur tumorale de moins de 4mm a été identifiée comme un facteur prédicteur de la survie spécifique et la survie à long terme dans plusieurs études, mais la PI a été identifiée comme facteur prédictif plus précis.[8]

2.3 Impact clinique de la PI

2.3.1 Rôle pronostique de la PI

Comme la langue est constituée uniquement de muscles intrinsèques et extrinsèques sous la surface épithéliale, aucun autre frein à l'extension tumorale n'existe une fois que l'invasion tumorale a dépassé la membrane basale. De plus, la langue possède un apport vasculaire et lymphatique plus riche que plusieurs autres sous-sites oncologiques de la tête et du cou. [9] Ainsi, plusieurs études ont récemment détaillé le rôle de la PI sur la présence de métastases occultes, la survie spécifique à la maladie et la survie globale à long terme [10-12].

Ce sont ces résultats qui ont justifié l'incorporation de la PI dans AJCC8. L'étude phare rétrospective de 3149 spécimens pathologiques de tumeurs linguales effectuée par Ebrahimi *et al.* a démontré que la PI de plus de 5mm est un facteur indépendant de réduction de la survie spécifique à la maladie, même lors d'ajustement par régression multivariée tenant compte des

catégories pT[†], pN, de la thérapie adjuvante, du délai de traitement, de l'âge, du sexe, de l'extention nodale extracapsulaire et du statut de la marge de résection, soit envahie ou non par le cancer.[13] Cela a permis de développer un modèle alternatif de *staging* basé sur la PI de plus de 5mm, de 5 à 10 mm, ou de plus de 10mm, tel qu'incorporé dans le AJCC8. (Voir *Table 1*, p.74)

Outre la survie spécifique à long terme, l'un des facteurs les plus importants pour guider la prise en charge thérapeutique est la présence de métastases ganglionnaires occultes. Tel que détaillé plus bas, le risque estimé de métastases ganglionnaires occultes peut constituer une indication de réaliser un évidement cervical électif.

Depuis l'adoption de l'AJCC8, plusieurs études ont évalué les valeurs optimales de PI permettant de prédire la présence de métastases cervicales nodales occultes. Toutefois, la valeur optimale est encore débattue par de multiples études rétrospectives et ne concorde pas toujours avec AJCC8. À titre d'exemple, alors que Tam *et al.* ont démontré qu'une PI de 7,25mm était la valeur limite la plus prédictive de présence de métastases occultes[14], Shinn *et al* ont plutôt proposé un intervalle critique de 2 à 6mm[11] alors que plusieurs autres études ont proposé 4mm comme limite significative.[15, 16] Une étude randomisée contrôlée par D'Cruz *et al,* elle, supporte l'indication d'évidement cervical électif chez les patients avec des carcinomes épidermoïdes de la cavité orale de PI supérieure à 3mm. [17]

2.3.2 Prise en charge thérapeutique

La prise en charge thérapeutique du carcinome épidermoïde de la langue mobile est d'abord chirurgicale. Le guide de pratique du National Comprehensive Cancer Network (NCCN) est le guide conventionnellement suivi au Canada pour la prise en charge oncologique des tumeurs de la langue. L'exérèse de la tumeur primaire par glossectomie partielle peut être combinée ou non à un évidement ganglionnaire cervical.[18]

[†] Catégorisation effectuée selon la 7eme édition du AJCC (AJCC7)

2.3.2.1 Glossectomie partielle

La glossectomie constitue le traitement primaire principal. Pour que celle-ci soit effectuée en marges saine, la distance minimale entre le tissu néoplasique envahissant et le tissu sain doit-être de 5mm lors de la mesure histopathologique finale.

Toutefois, la marge de résection minimale diffère selon l'envahissement tumoral. Une lésion Tis peut-être réséquée avec des marges plus rapprochées, évitant la résection inutile de tissu environnant sain. Par ailleurs, la résection de lésion de stade Tis peut être effectuée par vaporisation au laser.

2.3.2.2 Évidement cervical électif

Tel que discuté précédemment, bien que la valeur critique de PI justifiant l'évidement cervical électif soit débattue, le NCCN recommande de réaliser un évidement cervical ipsilatéral électif chez les patients présentant une PI supérieure à 4mm. L'évidement cervical électif n'est toutefois pas indiqué pour une PI inférieure à 2mm. Outre la PI, le seul autre paramètre inclus dans ce guide de pratique est le diamètre maximal de la tumeur, mesuré à la surface épithéliale linguale. Si une tumeur est de stade cT3N0M0 ou cT4N0M0, soit en raison de sa PI ou encore en raison de son diamètre maximal, l'évidement cervical électif est toujours indiqué. [18]

Ainsi, la mesure de la PI des néoplasies linguales avant leur résection chirurgicale est cruciale pour une planification de traitement optimale. À cet effet, plusieurs outils diagnostics peuvent être employés.

2.3.3 Mesure de la PI

2.3.3.1 Palpation

La palpation digitale est employée par les chirurgiens cervico-faciaux pour fournir une mesure approximative de la profondeur d'invasion. Il s'agit également de la méthode suggérée dans le guide NCCN.[18] Toutefois, malgré que la différence de quelques millimètres de profondeur ait un impact majeur sur la décision thérapeutique, aucune étude n'a encore établi de corrélation entre la PI mesurée à la palpation et la PI finale mesurée à l'évaluation histopathologique de la tumeur.

2.3.3.2 Biopsies

Actuellement, deux études[19, 20] ont évalué la fiabilité de la mesure PI sur des biopsies prises au sein de la tumeur. Ces deux études sont mentionnées dans les articles présentés ci-bas. Malheureusement, l'absence de standardisation dans la technique de biopsie confèrent à ces études une validité externe moindre.

2.3.3.3 Imagerie médicale

Quatre études ont jusqu'à présent établi des corrélations entre la PI mesurée sur le *computed tomography* (CT) scan préopératoire et la PI pathologique. Ces études sont répertoriées et discutées dans le premier article ci-bas. Comme le scan fait partie du bilan d'investigations nécessaires pour éliminer la présence de métastases cliniquement détectables dans les néoplasies linguales[18], une mesure fiable de la PI sur les images scanographiques serait un ajout intéressant pour améliorer la planification chirurgicale sans réaliser d'imagerie additionnelle.

L'imagerie par résonnance magnétique (IRM) est une alternative au CT scan qui confère une meilleure résolution aux tissus mous de la tête et du cou. La précision et la fiabilité de la mesure de PI sur l'IRM a été plus largement étudiée dans la littérature actuelle. Dans l'article présenté ci-bas, les 23 études ayant établi des coefficients de corrélation entre la PI à l'IRM et à la pathologie sont présentées et comparées.

L'échographie linguale est une autre modalité d'imagerie qui présente des avantages dans la l'évaluation de lésions linguales. Elle est rapide, peu coûteuse, et ne génère pas d'irradiation néfastes pour le patient [21]. De multiples études ont établi des corrélations entre l'épaisseur tumorale échographique et pathologique, mais peu d'études se sont concentrées sur la PI. Puisque la valeur prognostique de ces deux paramètres diffère, la précision de l'échographie linguale pour la mesure précise de la PI reste à démontrer.

2.3.4 Questions de recherche

À la lumière de la nouvelle classification du AJCC, des recommandations publiées par le NCCN, la mesure préopératoire précise de PI est nécessaire pour décider de la largeur des marges

de résection (1mm pour le Tis ou 5mm pour l'invasif), de la méthode de résection (vaporisation au laser pour le Tis, ou glossectomie chirurgicale pour l'invasif), et de l'indication d'effectuer un évidement cervical ou non (selon la PI supérieure ou inférieure à 4mm).

Devant la multiplicité des méthodes récemment étudiées pour mesurer la PI en préopératoire, deux questions se posent. Premièrement, quel est l'ensemble des méthodes actuellement décrites dans la littérature pour mesurer la PI en préopératoire, et comment celles-ci se comparent-elles les unes aux autres? Deuxièmement, si certaines méthodes sont actuellement moins bien étudiées, pouvons-nous en développer une qui permette de mieux répondre à ce besoin clinique?

Le premier article présenté vise donc à répertorier et comparer les diverses modalités décrites jusqu'à présent dans la littérature pour mesurer la PI en préopératoire. Le second article vise à mieux étudier l'utilité de la biopsie au poinçon pour déterminer la PI.

Chapitre 3 – Présentation des articles

Article 1 - Preoperative evaluation of depth of invasion in oral tongue squamous cell carcinoma: a systematic review and meta-analysis

Authors: Béatrice Voizard MD FRCSC[1,2,4], Michel Khoury MD[1,2], Nadim Saydy MD MSc[1,2], Kristoff Nelson MD FRCPC[1], Guillaume B. Cardin MSc[3], Laurent Létourneau-Guillon MD MSc FRCPC[1,2,3,] Abdelali Filali-Mouhim[3], Apostolos Christopoulos MD MSc FRCSC[1,2,3]

Affiliations:

[1]Faculty of Medicine, Université de Montréal, Montreal, Canada

[2]Division of Otolaryngology – Head and Neck Surgery, Department of Surgery, Université de Montréal, Montreal, Canada

[3]Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, Canada

[4] Centre Intégré Universitaire de Santé et de Services Sociaux (CIUSSS) du Nord-de-l'Île-de-Montréal, Montreal, Canada

Corresponding author:

Apostolos Christopoulos, MD MSc FRCSC
a.christopoulos@umontreal.ca
Associate Professor, Division of Otolaryngology-Head and Neck Surgery
Université de Montréal
Permanent address : Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM)
900 St-Denis St
Montreal- QC, Canada, H2X 0A9
Tel: (514) 890-8000 - Fax: (514) 412-7328

Abstract

The inclusion of depth of invasion (DOI) in the American Joint Committee on Cancer's staging system for oral cavity squamous cell carcinoma (SCC) has major clinical impacts. Recent studies have evaluated the reliability of imaging modalities and biopsy techniques to measure DOI preoperatively. The objective of this systematic review and meta-analysis was to comprehensively include all previously described methods to measure preoperative DOI in oral tongue SCC (OTSCC) and to compare their reliability. A systematic review was conducted on PubMed, Embase and Cochrane according to the PRISMA guidelines. Studies that evaluated the reliability of DOI measured on biopsy or imaging (rDOI) by comparing it to DOI on histopathology (pDOI) were included for extraction. A meta-analysis was conducted to obtain pooled correlation coefficients for each imaging modality. The pooled correlation coefficients between rDOI and pDOI were 0.86 ($CI_{95\%}$=[0.82–0.88]) and 0.80 ($CI_{95\%}$=[0.70-0.87]) for magnetic resonance imaging (MRI) studies and computed tomography (CT) studies, respectively. For ultrasound (US), the correlation coefficient could only be measured by including studies which measured not only DOI but also tumor thickness. It was 0.89 ($CI_{95\%}$= [0.82-0.94]). Overall, MRI is the better studied modality. It has a good reliability to measure preoperative rDOI in OTSCC. CT is less studied but appears to be less reliable. US cannot be compared to these imaging modality as it has been used more often to measure TT than DOI.

Highlights

- Depth of invasion is a prognostic factor in oral tongue squamous cell carcinoma
- Magnetic resonance has good reliability to measure depth of invasion
- Ultrasound is mainly studied to measure tumor thickness and not depth of invasion
- Computed tomography has a lower reliability to measure depth of invasion

Keywords

Head and Neck Cancer, Oral cancer, Tongue Neoplasms, Squamous Cell Carcinoma, Biopsy, Diagnostic Imaging, Neoplasm Staging, Neoplasm Invasiveness

Introduction

Among head and neck cancers, the majority of which are squamous cell carcinomas (SCC), approximately 25% arise in the oral cavity with the oral tongue representing the most commonly affected subsite[3]. Treatment of oral tongue SCC (OTSCC) may consist of curative surgery with or without adjuvant radiotherapy and chemotherapy.[18] In recent years, the prognostic value of new pathological features has had major impacts on treatment planning. Among them, tumor depth of invasion (DOI) was defined as the distance from the level of the epithelial basement membrane (BM) to the tumor invasive front[13]. DOI differs from tumor thickness (TT), which is defined as total thickness of the tumor and combines the invasive depth and the exophytic portion above the BM[22].

In recent studies, pathologic DOI (pDOI) has proven to be a strong predictor of occult lymph node metastasis as well as disease-specific survival[13]. In light of these studies, the eighth edition of the American Joint Committee on Cancer (AJCC8) staging manual now incorporates pDOI in its tumor, node, metastasis (TNM) staging system.[5] Further, many studies have proposed various pDOI cut-off values to stratify patients for elective node dissections in early OTSCC (staged T1 or T2, N0).[16] Despite its crucial prognostic value, preoperative and postoperative evaluation of DOI remains a challenge, especially in early OTSCC.[23]

Various imaging modalities have been studied to precisely measure DOI in the preoperative setting. Among these studies, magnetic resonance imaging (MRI), ultrasound (US) and computed tomography (CT) have been widely studied. Correlations between radiologic DOI (rDOI) and pDOI measured on the final histopathology after glossectomy vary greatly, as do imaging protocols, studied populations and inclusion criteria. Previously published systematic reviews and meta-analysis (SRMA) have selectively included MRI[24] or ultrasound[25, 26], but none has yet comprehensively included all imaging modalities and biopsy techniques. The aim of this study is

to compare the reliability of current methods described in the literature to adequately measure preoperative DOI in OTSCC.

Methods

We conducted a systematic literature review according to the 2020 Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement guidelines[1].

Search strategy

A comprehensive systematic search was conducted in PudMed (MEDLINE), Cochrane and Embase databases for articles published between 1946 and 2022. Search terms were formulated both in plain language and using the appropriate thesauri terms of each database. The complete search formulation is shown in Supplementary file 1. A manual search was conducted in the reference section of each included article and in three previously published systematic reviews on the subject[24-26]. After identifying records through database researching and removing duplicates, abstracts were screened for eligibility by two independent readers (authors BV and MK) and classified as “included”, “excluded” or “neither excluded nor included”. A tertiary, independent reader further classified the “neither excluded nor included” abstracts, deciding whether they would be included or not. Included full-text articles were read and analyzed, and further articles were excluded.

Inclusion criteria

All included studies described the use of an imaging modality or biopsy technique to measure preoperative TT/DOI[‡] on OTSCC. All included studies provided either a correlation coefficient or the raw data to calculate the correlation between preoperative TT/DOI and final TT/DOI on

[‡] For ease of reading, TT/DOI will be used in lieu of « TT or DOI ».

histopathology. Although the aim of this SRMA was to measure preoperative DOI, both studies on TT and DOI were first included to broaden the scope of this review.

Exclusion criteria

We excluded all studies without a histological TT/DOI measurement as a gold standard. We also excluded all studies which only measured TT/DOI *in vitro* after the glossectomy or evaluated TT/DOI intraoperatively, as vigorous tongue retraction or other techniques done under anesthesia would not be feasible on an awake patient[27], diminishing the external validity of this review. Studies that did not provide a correlation coefficient but provided only Bland-Altman plots were excluded. Any study including cancers other than SCC, as well as animal studies were excluded. Studies measuring the correlation between immunohistochemical markers and TT/DOI were excluded[28]. Finally, non-peer reviewed publications, abstracts, scientific posters, podium presentations, letters to authors, responses to other articles and case reports were also excluded.

Some studies included both tongue tumors and tumors arising in other oral cavity subsites. If any variable such as TT/DOI values were presented as a cohort mean value and not categorized by subsites, the value was excluded from the analysis to ensure that alternative pathologies beyond the scope of this review were not included.

Qualitative appraisal of the studies

After full text assessment, one specialized neuro-radiologist (author KN) reviewed every study. Studies that used outdated MRI, CT or US technology or whose measurement methods were irreproducible were also excluded after this qualitative assessment.[29-45]

Data extraction and analysis

The following data were extracted from the included studies: number of included patients, age, sex distribution, study type, whether TT or DOI was measured, inclusion and exclusion criteria of

each study and T-stage distribution when available. The MRI sequences, the US probe and technique, and the CT planes were also recorded. When a study reported separate correlation coefficients for a single patient population, they were all extracted and included in the meta-analysis as different point on the Forest plots.

Statistics

Meta-analyses were carried out using the R[46] (R Core Team 2019) package meta[47]. The Hedges-Olkin method[48] was first used to calculate the weighted summary correlation coefficient under the fixed effects model, using Fisher's z transformation of the correlation coefficients. Next, the heterogeneity statistic was incorporated to calculate the summary correlation coefficient along with its prediction intervals[49] under the random effects model. Prediction intervals are a more informative quantity for reporting variability in meta-analysis studies. They provide a range into which the effect of future studies is expected to fall based on present evidence.

To assess the robustness of our pooled results, we performed outlier analyses to detect studies with extreme effect sizes and estimate their effect on the pooled correlation coefficients and prediction intervals. The assumption of homogeneity was assessed by the Cochran Q test[50], and the degree of inconsistency across studies was calculated using Higgins & Thompson's I^2 statistic[49]. I^2 heterogeneity describes the percentage of total variation across studies that is due to studies' heterogeneity rather than sampling error. Values of 25, 50, and 75% indicate low, moderate, and substantial heterogeneity respectively.[51] As between-study heterogeneity could be caused by one or more studies with extreme effect sizes that may distort the pooled effect estimate, outliers detection analysis[52] was performed using the function « find.outliers() », in the R package dmetar. This package implements an outlier removal algorithm searching for outlier studies, excluding them, and then recalculates new pooled correlation coefficient, Cochran Q test, I^2 statistic and prediction interval. The R version 4.1.3 was used for the analysis.

Results

Search strategy and article selection

Figure 1 is a flow diagram representing the number of records identified, included, and excluded, and reason for exclusion when applicable. Among the 474 abstracts screened for eligibility, 42 disagreements were resolved by the third reader.

Study characteristics

In total, 36 studies were included in this systematic review, totaling 3226 TT/DOI measurements. Among the included articles, there were 23 studies on MRI, 13 on US, 4 on CT and 2 on biopsy techniques. Four included studies[53-56] compared two or more imaging modalities, as detailed in Figure 2.

General characteristics of studies focusing on imaging modalities are reported in Table 1. Most studies came from Asia. Included studies were published between 1997 and 2022. AJCC8 was published in 2017, but its implementation was delayed to January 2018. Therefore, studies were stratified as published either before or after 2018. A majority (58%) of included imaging studies were published after 2018. Among studies published after AJCC8 implementation, 81% reported DOI rather than TT measurements.

Table 2 summarizes the extracted data for all the included studies. Inclusion and exclusion criteria varied widely, making the pooled study population heterogeneous. Insufficient data was reported for patient sex distribution and age range. Included T stages and staging systems varied depending on the publication date of each study. Early OTSCC (stages T1-2) was the main subject of included studies for all imaging modalities, but many studies also included more advanced

tumors (stages T3-4). As many studies did not report sufficient data and staging systems evolved over time, result stratification according to T stages was not possible.

Exclusion criteria varied widely. Some studies[57, 58] included undetectable lesions and gave them a value of rDOI=0mm, whereas most excluded undetectable lesions. Some other studies[54] excluded rDOI<3mm because they were deemed undetectable on MRI, whereas other studies measured and included these values. Some studies excluded patients who had a biopsy before imaging[59, 60] as the resulting inflammation was expected to distort the tissue and change the rDOI. Others excluded patients for whom the delay between imaging and surgery exceeded 2[61] to 7[62] or 8 weeks[60], as the tumor growth between measurements could cofound the correlation. Many studies excluded patients who received neoadjuvant treatments. Not all studies detailed their exclusion criteria fully. Exclusion criteria for each included imaging study is provided in Supplementary files 2, 3 and 4.

Measurement of DOI on biopsy

Since there were only two studies focusing on the use of biopsy to calculate preoperative DOI, it was impossible to calculate intraclass correlation coefficients between two studies. Almangush *et al*.[19] used incisional biopsy to evaluate DOI, whereas Moore *et al.*[63] did not describe their biopsy method. Therefore, no qualitative comparison can be made between these studies.

Correlation between rTT, rDOI and pDOI

The random-effects meta-analysis models were chosen because studies were heterogeneous, as demonstrated by high I^2 values (Figures 3, 4 and 5).

MRI

Among included studies on MRI, 22% measured TT as opposed to DOI (78%). (Table 1) Most (91%) MRI studies focused on early OTSCC (T1-T2). The most used MRI sequences were contrast enhanced T1. (Table 2)

The first correlations included all MRI studies that measured either rTT or rDOI (Figure 3.a). The pooled correlation coefficient between rTT/rDOI vs pTT/pDOI was 0.84 ($CI_{95\%}$=[0.80–0.87]). The prediction interval for these pooled studies was [0.54–0.95]. Therefore, if another study correlating rTT/rDOI on MRI to pTT/pDOI was conducted, the correlation coefficient would fall within that range.

When calculating a pooled correlation coefficient including MRI studies which measured only rDOI, and excluding the ones measuring rTT, the correlation and prediction intervals were equivalent. (Figure 3.b) Therefore, MRI's ability to measure either rTT or rDOI did not differ.

When removing outlier studies, in order to lower between-study heterogeneity (optimising the I^2 statistic), the correlation between rDOI and pDOI was 0.86 ($CI_{95\%}$=[0.82–0.88]) and the prediction interval was narrowed to [0.67–0.94] (Figure 3.c).

US

Among included studies on US, 79% measured TT as opposed to DOI (21%). Because only few studies measured DOI, no meta-analysis was conducted on this subgroup.

As presented in Figure 4, the pooled correlation coefficient for the US studies measuring rTT/rDOI was 0.89 ($CI_{95\%}$= [0.82-0.94]). This makes it the highest correlation coefficient among all imaging modalities. Of note, the I^2 heterogeneity was also the highest among all modalities at 94%. The

prediction interval was therefore wide ([0.26-0.99]). High variability among US techniques is also described in Table 2. Four studies did not describe their techniques. Some studies used a water-filled glove to reduce the risk of compressing the tumor, whereas others used only sterile gel. Insufficient information was provided to compare US techniques among them.

CT

All included CT studies measured rDOI. None measured rTT. The coronal plane was most commonly chosen to measure rDOI.

The pooled correlation coefficient was 0.85 ($CI_{95\%}$=[0.73–0.92]). (Figure 5.a) The prediction interval was wide at [0.24-0.98] and the I^2 heterogeneity between studies was substantial (I^2= 77%).

As presented in Figure 5.b, when excluding outlier studies to reduce heterogeneity, the prediction interval was narrowed to [0.51-0.93], but the pooled correlation coefficient was lower (0.80 ($CI_{95\%}$=[0.70-0.87])) . The I^2 heterogeneity remained higher than moderate, at 69%. Therefore, it appears that CT studies are heterogeneous, and the correlation between rDOI and pDOI on CT is lower than on MRI.

Discussion

To our knowledge, this is the first comprehensive SRMA that includes all the imaging and biopsy modalities studied to determine preoperative DOI in OTSCC. Unlike previous SRMAs which had focused mainly on a single imaging modality, this review aims to apply the same systematic search criteria and meta-analysis methods to all imaging modalities with the objective of adequately comparing among them. Nonetheless, this comparison is limited by the fact that most recent studies measure DOI and not TT, which is in accordance with AJCC8.

MRI

Because of the relatively few numbers of CT and US studies focusing on rDOI rather than rTT, comparing the reliability of CT, US and MRI to precisely measure rDOI presents a challenge. Overall, the robustness of the pooled data is stronger in MRI studies on rDOI after the exclusion of outlier studies, with a moderate I^2 heterogeneity of 58%. (Figure 5.c)

A recent SRMA by Li et al.[24] compared 9 studies on MRI's reliability to measure preoperative TT and DOI in OTSCC. They obtained an overall intraclass coefficient (ICC) of 0.89 ($CI_{95\%}$= [0.837-0.895]), which is comparable to our results of 0.84 ($CI_{95\%}$=[0.80–0.87]). Of note, the I^2 heterogeneity of 30.9% found in their study was lower than ours, and they included fewer studies. The present SRMA differs since we stratified studies measuring only rDOI. Measuring rDOI requires further image manipulation compared to rTT. It is calculated by tracing a reference line that connects the normal mucosal junction on both sides of the tumor. Then, DOI is measured perpendicular to this line, down to the most invasive point of the tumor into the normal tongue stroma.[53, 64] Although it might be expected that such image manipulation would result in a higher margin for error and diminish the reliability of MRI, one study by Mair *et al.* showed a very low inter-observer variability between two experienced head and neck radiologists measuring tumors independently. A correlation coefficient of 0.965 ($p<0.001$) was obtained between them.[65] Our correlation coefficient of 0.86 ($CI_{95\%}$=[0.82–0.88]) is still in keeping with the previous findings of Li *et al.* showing that MRI has a good reliability for both rTT and rDOI.

The absence of standardised MRI protocols may hinder the external validity of these findings. The distribution of T-stages and the location of tumors on the dorsal, lateral, or ventral surfaces of the mobile tongue varied across studies as listed in Table 2 and Supplementary file 2. All these factors may have a significant impact on the optimal MRI sequence chosen by specialised head and neck radiologists to measure rDOI. Overall, the provided information was insufficient to

conduct a clinically significant subgroup analysis comparing the correlations between different MRI sequences.

US

In the present review, US was the second most widely studied imaging modality. In a recent SRMA by Klein Nulent *et al.* the tongue was the most commonly studied subsite of the oral cavity among twelve included studies. The pooled individual participant data Pearson's correlation for rTT and pTT was 0.88. Interestingly, after stratifying lesions according to clinical T stage, they showed a better correlation with increasing stage.[26] Although this SRMA differs from ours as it included US performed intraoperatively, the results are still in keeping with our pooled correlation coefficient of 0.89 ($CI_{95\%}$= [0.82-0.94]). In another SRMA by Tarabichi *et al.,* ten studies were pooled using a random effects model and resulted in a correlation coefficient of 0.95 ($CI_{95\%}$= [0.89-0.98]), which is slightly higher than our results. The I^2 heterogeneity was measured at 94.4%, which is equivalent to the heterogeneity found in our included studies. In both SRMAs, most studies on US often used the terms rTT and rDOI interchangeably since many were published before data on DOI became widespread. Our pooled results include TT and DOI and can effectively be compared with these two other SRMAs. Overall, despite showing the highest heterogeneity of all imaging modalities, US has a good reliability as a predictor of pTT/pDOI.

US is subject to inherent variations due to inter-reader reliability and variation in techniques. Although this is beyond the scope of this review, it may account for some of the variability observed in US, where the I^2 was highest compared to CT and MRI.

CT

This is the first SRMA on CT's reliability to measure rDOI. Although only few studies were found, they still comprised of 157 rDOI measurements. The pooled correlation coefficient of 0.80

($CI_{95\%}$=[0.70-0.87]) indicates a moderate to good reliability[66]. In practice, CT is a rapid, widely available, imaging modality that is often used to screen for lymph node metastasis. Therefore, contrary to MRI or US, CT is often obtained systematically in oral cavity invasive SCC cases. In the future, inclusion of systematic rDOI measurement on CT could represent an easy, low-cost clinical aid to make better treatment planning without additional imaging. Further studies are required to lower the heterogeneity of the results.

Limitations

There are many limitations to this systematic review and meta-analysis. The included population was heterogeneous. The evolution of the AJCC TNM staging systems over time prevented adequate data stratification of the correlation coefficients according to T-stage. While some studies provided raw data, most did not. Therefore, even studies which only calculated correlation coefficients on T1-T2 or on T4a tumors were included. It is expected that certain rDOI measurement techniques may be optimised for some T-stages as suggested by Klein Nulent *et al*.[26], but the present study cannot draw such conclusions.

Only a few studies[59, 60] excluded patients who had a biopsy before the measurement of rDOI, while some others stratified patients according to timing of biopsy.[21, 54] Inflammation, bleeding, and scarring may distort the tissue and confound the measurement. Yesuratnam *et al*. showed that performing biopsy before either US or MRI could increase the mean difference between rTT and pTT by 0.63mm and 2.44mm, respectively.[21] The clinical impact of such a small difference in rTT/rDOI measurement remains understudied. For instance, Howe *et al*. compared pre-biopsy and post-biopsy MRI's ability to adequately determine T-stage in OTSCC. They showed no significant difference in the upstaging or downstaging of tumors after histopathologic analysis.[67] Therefore, although biopsy and imaging timing may confound the correlation between DOI measurements, it may not prevent adequate staging of the tumor in the preoperative setting. Irrespective of T stage, a difference of >2 mm in DOI could have a major impact on the decision to perform an elective cervical neck dissection for early (T1-T2, N0) OTSCC.

During the process of formalin fixation, sectioning and final preparation, any tumor specimen may be slightly distorted, providing yet another potential confounding factor between real tumor DOI, rDOI and pDOI. One study by Harada *et al.* calculated a formula to account for such tumor shrinkage between excision and histopathologic measurement.[54] They did not, however, compare separate correlations between rDOI and non-adjusted pDOI vs pDOI adjusted for shrinkage. Therefore, little conclusions can be made as to whether tumor shrinkage effectively confounds the correlation.

Clinical feasibility

Many clinical factors may hinder the feasibility of each imaging modality. US studies cited gag reflex, pain and movement as limiting factors during the procedure.[68] On MRI, patient movement has been cited as a potential problem[54, 59], while dental artifacts may impair the measurements on CT[69]. Rapid access to any given imaging modality is crucial if a surgeon plans to rely on rDOI for surgical planning. Delay between imaging and surgery is a major potential confounding factor in most of the included studies. As previously mentioned, only a few studies[60, 61, 70] excluded patients with long imaging to surgery intervals to minimise the impact of tumor growth between each measurement.

Finally, some studies established direct correlations between rDOI and clinical prognostic parameters, irrespective of the correlation between rDOI and pDOI. For instance, Natori *et al.* demonstrated that TT on US was a significant predictor of the risk of neck metastasis in early (T1-2, N0) OTSCC.[71] Xu *et al*. showed that an MRI-determined DOI of >7.5mm was a significant predictor of both occult LN metastasis and loco-regional control in OTSCC, independent of lympho-vascular invasion, pathologic tumor grade and pDOI in T1N0 OTSCC.[72] These correlations between clinical parameters and rDOI are what could allow surgeons to reliably use rDOI to determine preoperatively whether or not to perform an elective cervical neck dissection

for a given tumor, without waiting for a final pathology report. In the future, we suggest that all studies correlating rDOI and pDOI should also include direct correlations between rDOI and clinical prognostic parameters.

Conclusion

To our knowledge, this is the first SRMA comparing MRI, US and CT's reliability to measure preoperative DOI in OTSCC. MRI was the better studied imaging modality, and it has a good reliability with a pooled correlation coefficient of 0.86 between rDOI and pDOI. In comparison, studies on US focused more on measuring rTT than rDOI and have a very high heterogeneity. Both our results on US and MRI are in keeping with previously published SRMAs. CT scan has been used exclusively to measure rDOI and shows a slightly inferior reliability, but more studies on CT are needed to validate these findings. This review also highlights that the measurement of DOI on biopsy specimens remains largely understudied and may be an interesting area of future research. Finally, subsequent studies on this subject should directly correlate rDOI with clinical outcome measures, as this could prove whether the use of a given imaging modality should effectively alter preoperative treatment planning, especially in early OTSCC where DOI dictates the need for an elective cervical neck dissection.

Declarations

Availability of data and material

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the Dr Azar-Angelil Research Chair in Head and Neck Oncology of the Université de Montréal (AC).

Author's contribution

Béatrice Voizard MD FRCSC : study design, data collection, manuscript redaction, manuscript revision

Michel Khoury MD : study design, data collection

Nadim Saydy MD : data collection, manuscript revision

Kristoff Nelson MD, FRCPC : study design, data collection, manuscript revision

Guillaume B. Cardin : manuscript redaction, manuscript revision

Laurent Létourneau-Guillon MD MSc FRCPC : manuscript redaction, manuscript revision

Ali Filali : statistics, data collection, manuscript redaction

Apostolos Christopoulos MD MSc FRCSC : manuscript redaction, manuscript revision

Figure captions

Figure 1: PRISMA Flow Diagram of the study selection process

Figure 2: Venn diagram of included DOI-measurement studies

Figure 3: Forest plots showing the effect size of the correlation between pDOI and MRI-derived rTT/rDOI.

Caption: Estimates are at the center of the boxes and drawn in proportion to ES. Lines indicate 95% CIs. Diamonds shows the pooled effect size at its center and 95% CI at its horizontal points. The pooled correlation coefficient with 95% CI is given both for the fixed and random effect models along with I^2 statistic and the prediction interval.

Figure 4: Forest plots showing the correlation between pDOI and US-derived rTT and rDOI.

Caption: Estimates are at the center of the boxes and drawn in proportion to ES. Lines indicate 95% CIs. Diamonds shows the pooled effect size at its center and 95% CI at its horizontal points. The pooled correlation coefficient with 95% CI is given both for the fixed and random effect models along with I^2 statistic and the prediction interval.

Figure 5: Forest plots showing the correlation between pDOI and CT-derived rDOI.

Caption: Estimates are at the center of the boxes and drawn in proportion to ES. Lines indicate 95% CIs. Diamonds shows the pooled effect size at its center and 95% CI at its horizontal points. The pooled correlation coefficient with 95% CI is given both for the fixed and random effect models along with I^2 statistic and the prediction interval.

Figure 1. – PRISMA Flow Diagram of the study selection process

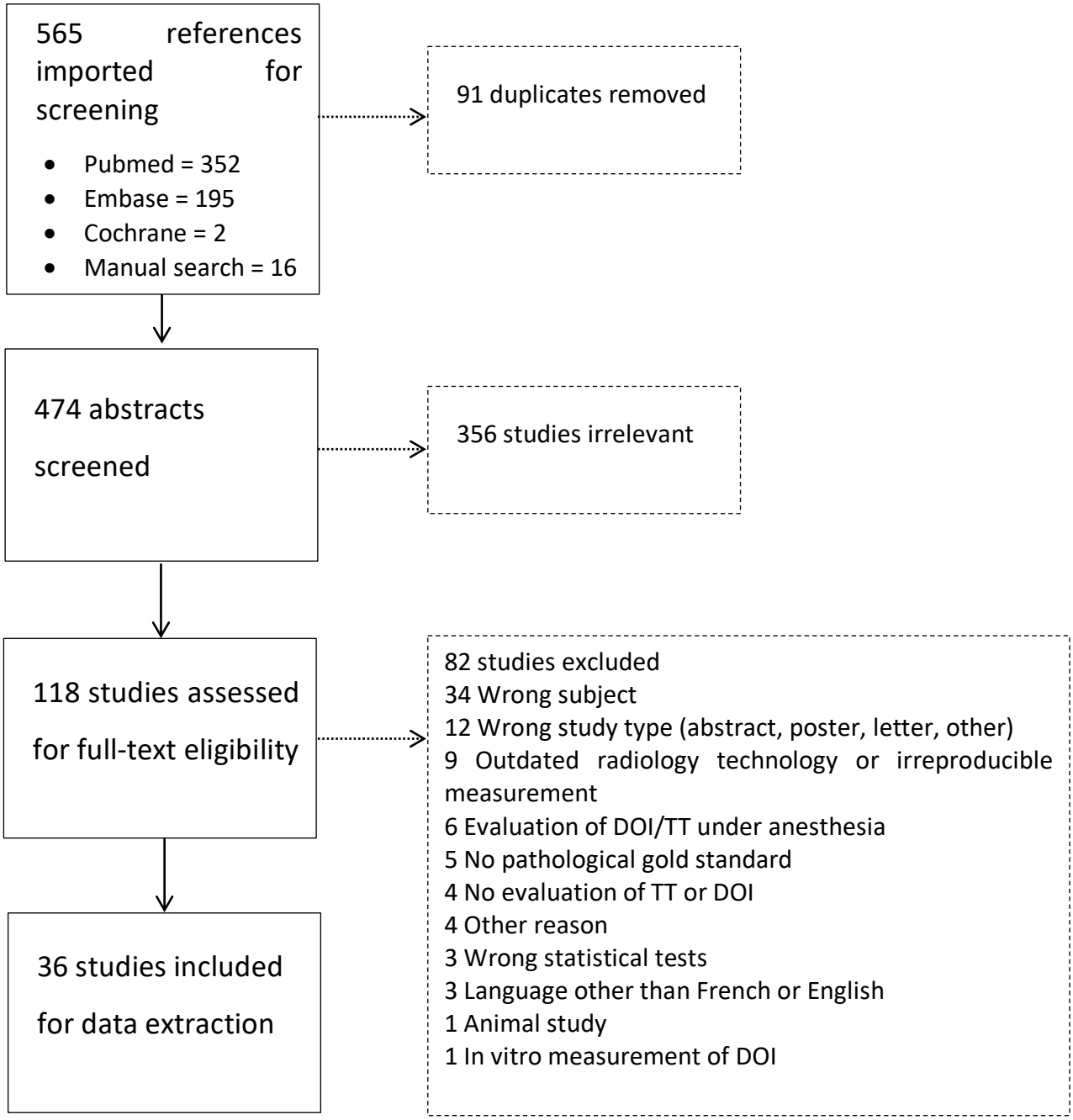
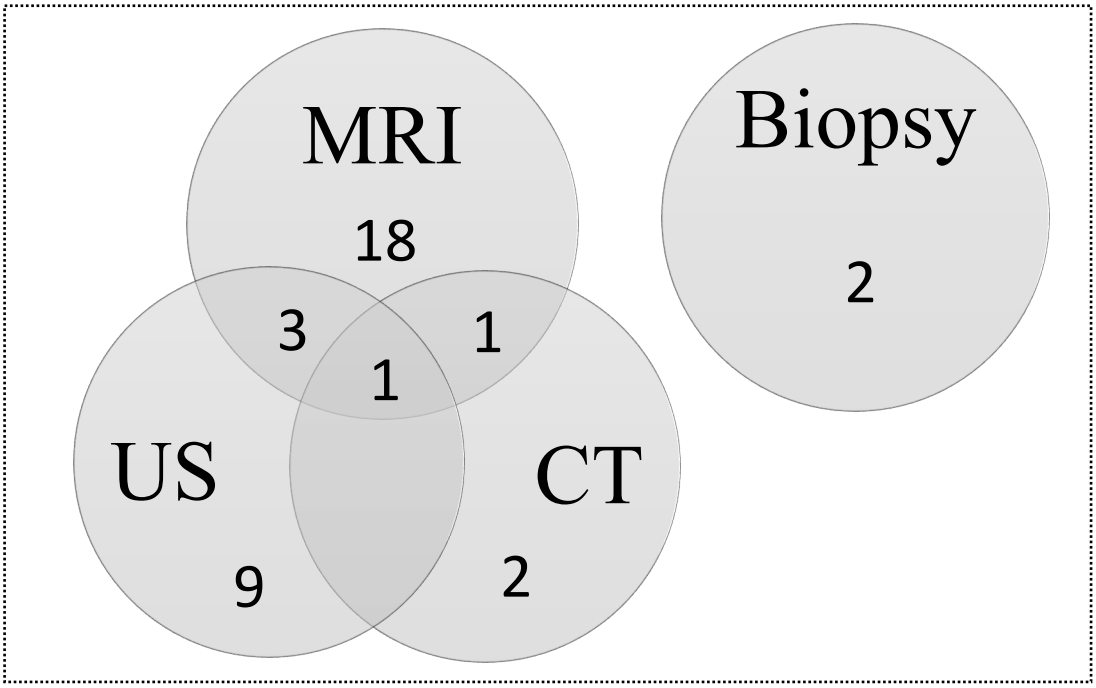


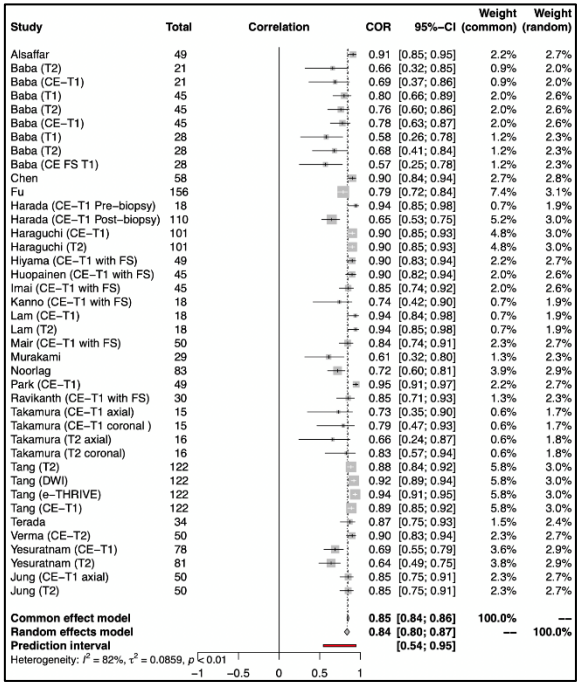
Figure 2. – Venn diagram of included DOI-measurement studies



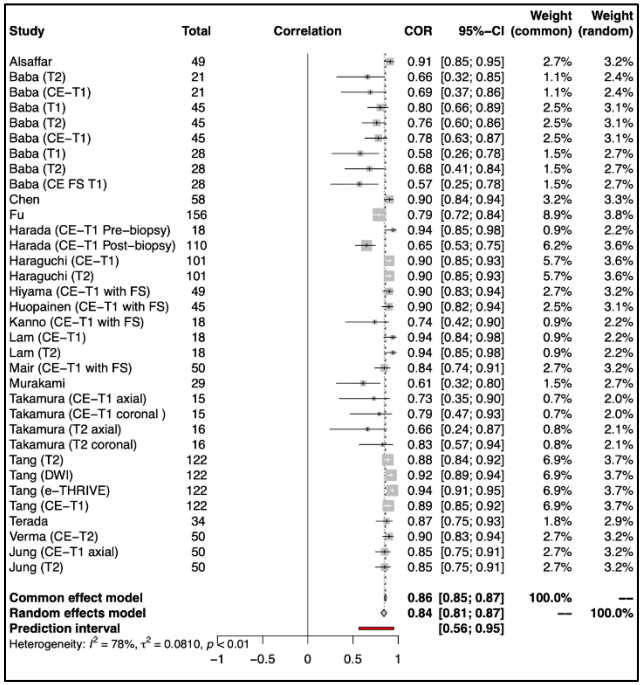
Of note, for Kurokawa[73] *et al.* we excluded the MRI and CT values because of outdated imaging technology but kept the US values.

Figure 3. – Forest plots showing the effect size of the correlation between pDOI and MRI-derived rTT/rDOI.

(a) All included MRI studies measuring rTT and rDOI



(b) MRI studies measuring rDOI



(c) MRI studies measuring rDOI (optimised I2)

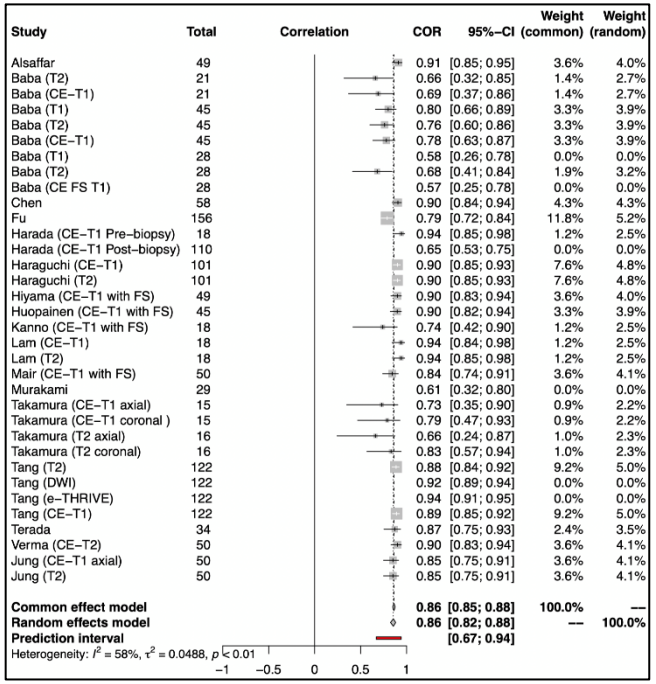


Figure 4. – Forest plots showing the correlation between pDOI and US-derived rTT and rDOI.

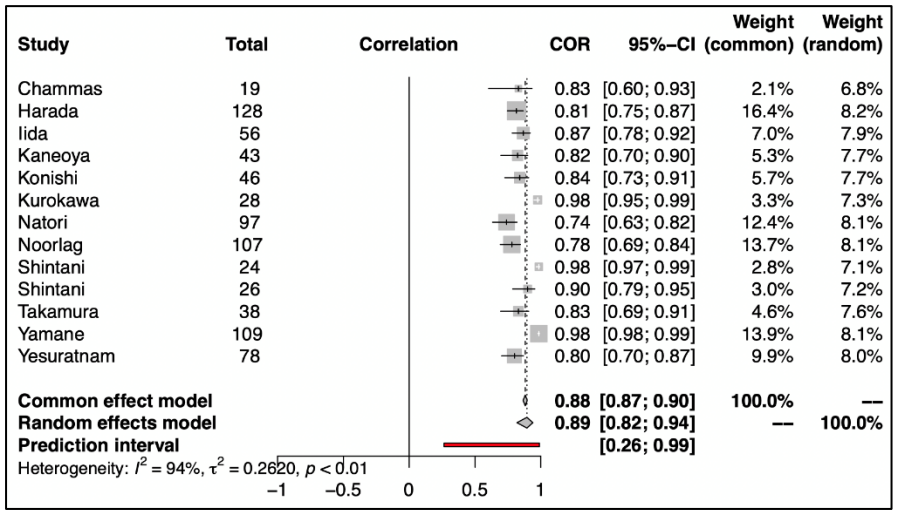
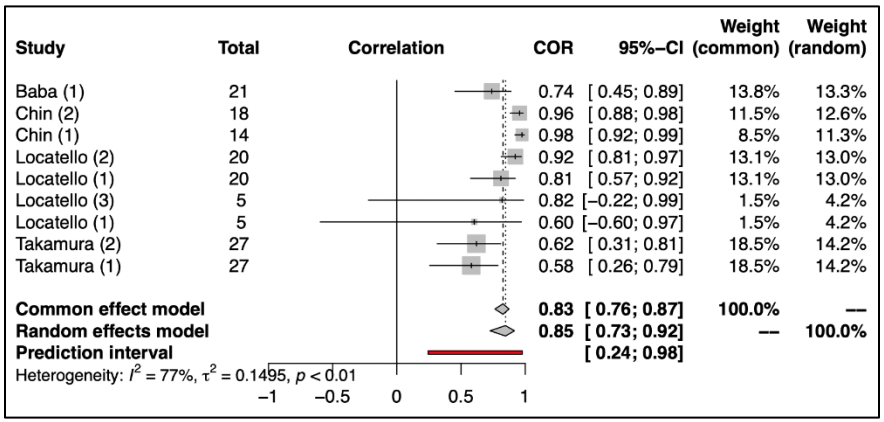


Figure 5. – Forest plots showing the correlation between pDOI and CT-derived rDOI.

(a) All included CT Studies measuring rDOI



(b) All included CT Studies measuring rDOI (optimised I^2)

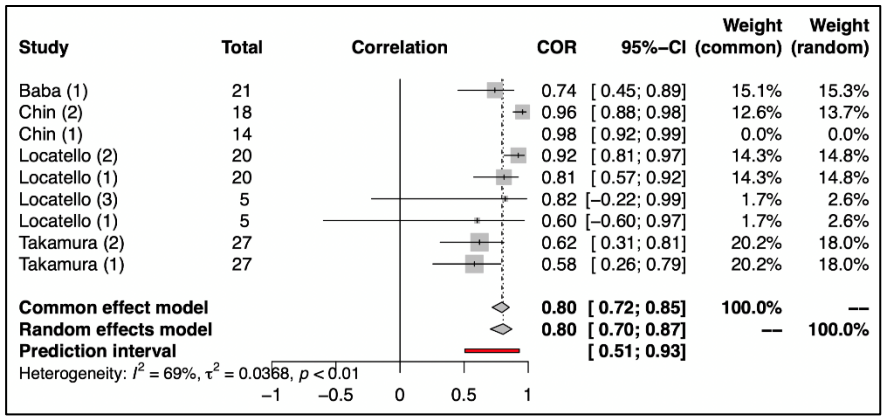


Table 1. – Pooled characteristics of studies on imaging modalities

Imaging modality		**MRI**	**US**	**CT**
Included studies	Number of studies	23	13	4
Patient population	Total TT + DOI measurements	2152	799	157
	Range number patients per study	15-156	19-128	5-27
Origin	Asia	18	10	3
	Europe	3	1	1
	Other	2	2	0
Publication year	1997-2017]	7	8	0
	[2018- present	16	5	4
Measurement	TT	22%	69%	0%
	DOI	78%	31%	100%
Included T stages	T1 - T2	91%	85%	75%
	T3 +/- T4	65%	54%	50%
	NA	4%	15%	25%

*NA : not available

Table 2. – Summary of included studies

Author	Year	Country	#Patients	Study type	Measure	T-stage included	MRI sequence	Correlation coefficient	p-value or CI	Coefficient
						MRI				
Alsaffar	2016	Canada	49	P	DOI	1,2,3,4 AJCC7	NA	0.91	p <0.001	Spearman
Baba	2018	Japan	45	R	DOI	1,2 AJCC7	T1	0.8	p <0.001	Spearman
							T2	0.76	p <0.001	
							CE-T1	0.78	p <0.001	
Baba	2019	Japan	28	R	DOI	1,2 AJCC7	T1	0.58	p <0.001	Spearman
							T2	0.68	p <0.001	
							CE FS T1	0.57	p <0.001	
Baba	2021	Japan	21	R	DOI	1,2,3 AJCC8	T2	0.66	p <0.001	Spearman
							CE-T1	0.69	p <0.001	
Chen	2012	Taiwan	58	R	DOI	T4a AJCC6	T1, T2, T2-FS or CE-T1	0.905	<0.001	Pearson
Fu	2021	China	156	R	DOI	1,2,3 AJCC8	NA	0.79	<0.01	Pearson
Harada	2021	Japan	18	R	DOI	NA	CE-T1 Pre-biopsy	0.944	<0.001	Pearson
			110	R	DOI	NA	CE-T1 Post-biopsy	0.649	<0.001	
Haraguchi	2021	Japan	101	R	DOI	1,2,3,4 AJCC8	CE-T1	0.9	<0.05	Pearson
							T2	0.9	<0.05	
Hiyama	2021	Japan	49	NA	DOI	1,2,3,4 AJCC8	CE-T1 with FS	0.84-0.90	<0.05	Spearman
Huopainen	2021	Finland	45	R	DOI	1,2,3 IUCC8	CE-T1 with FS	0.898	<0.001	Pearson
Imai	2017	Japan	45	R	TT	T1N0 NA	CE-T1 with FS	0.85	<0.001	Pearson
Jung	2009	Korea	50	R	DOI	1,2 AJCC6	CE-T1	0.851	< 0.001	Pearson
							T2	0.813	< 0.001	
Kanno	2020	Japan	18	R	DOI	1,2,3 AJCC8	CE-T1 with FS	0.74	NA	Pearson
Lam	2004	Hong Kong	18	P	DOI	1,2,3 NA	CE-T1	0.939	<0.0005	Pearson
							T2	0.941	<0.0005	
Mair	2021	United Kingdom	50	R	DOI	1,2,3,4 AJCC8	CE-T1 with FS	0.844	<0.001	Pearson
Noorlag	2020	Netherlands	83	R	TT	1,2, AJCC7	STIR or CE-T1 with FS	0.72	< 0.0001	Pearson
Park	2011	South Korea	49	R	TT	1,2,3,4 AJCC6	CE-T1	0.949	< 0.001	Pearson
Ravikanth	2020	India	30	P	TT	1,2,3,4 AJCC8	CE-T1 with FS	0.851	< 0.001	Pearson
Takamura	2022	Japan	15	R	DOI	1,2 AJCC8	CE-T1 axial	0.73	< 0.05	Spearman
							CE-T1 coronal	0.79	< 0.05	
			16				T2 axial	0.66	< 0.05	
							T2 coronal	0.83	< 0.05	
Tang	2022	China	122	R	DOI	1,2,3 AJCC8	T2	0.885	< 0.001	Pearson
							DWI	0.92	< 0.001	
							e-THRIVE	0.936	< 0.001	
Tang	2022	China	122	R	DOI	1,2,3 AJCC8	CE-T1	0.89	< 0.001	Pearson
Terada	2020	Japan	34	R	DOI	1,2 AJCC7	NA	0.87	< 0.05	Pearson

Verma	2019	India	50	P	DOI	1,2,3 AJCC7	CE-T2	0.9	< 0.001	Pearson
Yesuratnam	2014	Australia	78	P	TT	1,2,3,4 AJCC6	CE-T1	0.69	95% CI [0.55–0.79]	Pearson
			81				T2	0.64	95% CI [0.49–0.75]	
						US				
Chammas	2011	Brazil	19	P	TT	1,2,3,4 AJCC7	Gel, WFS	0.83	<0.01	Pearson
Harada	2021	Japan	128	R	DOI	NA	Gel, xylocaine	0.815	<0.001	Pearson
Iida	2018	Japan	56	R	DOI	NA	WFS, no gel	0.867	<0.001	Spearman
Kaneoya	2009	Japan	43	NA	TT	1,2,3 UICC	Gel, xylocaine	0.824	<0.0001	Pearson
Konishi	2021	Japan	46	R	TT	1,2 UICC8	Gel, WFS	0.84	<0.0001	Pearson
Kurokawa	2005	Japan	28	NA	TT	1,2,3,4 UICC4	NA	0.976	<0.0001	Spearman
Natori	2008	Japan	97	R	TT	1,2,3,4 UICC2002	NA	0.74	< 0.001	Regression
Noorlag	2020	Netherlands	107	R	TT	1,2, AJCC7	NA	0.78	< 0.0001	Pearson
Shintani	1997	Japan	24	R	TT	1,2,3,4, NA	No Gel	0.985	< 0.001	Pearson
Shintani	2001	Japan	26	P	TT	1,2,3,4 UICC5	Gel	0.90**	<0.001	Pearson
Takamura	2022	Japan	38	R	DOI	1,2 AJCC8	NA	0.83	<0.05	Spearman
Yamane	2007	Japn	109	P	TT	1,2 AJCC5	Gel	0.985	<0.001	Pearson
Yesuratnam	2014	Australia	78	P	TT	1,2,3,4 AJCC6	Gel	0.8	95% CI 0.71–0.87	Pearson
						CT				
Baba	2021	Japan	21	DOI	R	1,2,3 AJCC8	Coronal	0.74	<0.001	Spearman
Chin	2021	Malaysia	18	DOI	R	1,2,3,4 AJCC8	Axial	0.956	<0.001	NA
			14				Coronal	0.975	<0.001	
Locatello	2020	Italy	20	DOI	R	NA	Axial	0.923	<0.01	NA
			20				Coronal	0.811	<0.01	
			5				Sagittal	0.821	0.089	
			5				Coronal	0.603	0.285	
Takamura	2022	Japan	27	DOI	R	1,2 AJCC8	Axial	0.62	<0.05	Spearman

*NA : not available. R : retrospective. P : prospective. CE ; contrast-enhanced. FS ; Fat saturation. WFS; water-filled sheath. **calculated based on provided raw data

Supplementary file 1: Detailed research strategy

All searches were done on February 2, 2022.

Embase search strategy

Emtree search

1 depth.mp. (358499)

2 tumor invasion/ (56719)

3 tongue.mp. or tongue/ (70445)

4 oral tongue.mp. (1693)

5 cancer.mp. or malignant neoplasm/ (4016282)

6 squamous cell carcinoma/ (125293)

7 imaging.mp. or imaging/ (2094484)

8 mri.mp. or nuclear magnetic resonance imaging/ (980073)

9 CT scan.mp. or x-ray computed tomography/ (171304)

10 ultrasound/ (201044)

11 biopsy.mp. or biopsy/ (951614)

12 3 or 4 (70445)

13 7 or 8 or 9 or 10 or 11 (3119345)

14 5 or 6 (4046998)

15 depths.mp. (27716)

16 1 or 2 or 15 (425819)

17 12 and 13 and 14 and 16 (195)

- Filters applied:

French language

English language

1974 to 2022 February 02

Pubmed search strategy

Plain text search

((depth of invasion) AND (oral tongue)) AND ((cancer) or (squamous cell carcinoma)) AND ((imaging) OR (biopsy))

MeSH terms search

("Tongue Neoplasms"[Mesh] AND "Carcinoma, Squamous Cell"[Mesh]) AND ("Biopsy"[Mesh] OR "Diagnostic Imaging"[Mesh] OR "diagnostic imaging" [Subheading]) AND ("Neoplasm Staging"[Mesh] OR "Neoplasm Invasiveness"[Mesh])

- Filters applied:

French language
English language
1946 to 2022 February 02

Cochrane Databases search strategy

Plain text search

((depth of invasion) AND (oral tongue)) AND ((cancer) or (squamous cell carcinoma)) AND ((imaging) OR (biopsy))

Filters applied:

Word variations were not searched

Included databases: Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews

Supplementary file 2: Exclusion criteria of included MRI studies

Author	Year	Exclusion criteria
Alsaffar	2016	MRI scan that was reviewed by the surgeon prior to the clinical exam Carcinoma in situ Previous excisional biopsy Previous head and neck radiation or chemoradiation
Baba	2018	Dorsal, tip and ventral tongue lesions. Undetectable lesions on MRI excluded (instead of getting a DOI=0mm value). N+ necks
Baba	2019	Undetectable lesions on MRI excluded (instead of getting a DOI=0mm value). Dorsal, tip and ventral tongue lesions Extrinsic tongue muscle invasion Ventral and dorsal lesions N+ neck
Baba	2021	Dorsal, tip and ventral tongue lesions
Chen	2012	M+ Artifacts on MRI that made the measurement of TT impossible Loss to follow up
Fu	2021	Multiple lesions in oral cavity Previous surgery, radiation or other treatment in oral cavity or neck MRI performed >50days before the operation
Harada	2021	Treated in another center Peroperative neoadjuvant treatment Very exophytic (>3mm) or remarkably ulcerated (>3mm) tumor pDOI < 3mm because undetectable on MRI
Haraguchi	2021	CIS Neoadjuvant chemotherapy or radiotherapy
Hiyama	2021	Previously operated CIS Lesions not visible excluded from coefficient calculation
Huopainen	2021	Previous head and neck chemotherapy or radiotherapy
Imai	2017	Prior treatment Adjuvant treatment Positive margins
Jung	2009	M+ Ventral tongue Tumor extending to the floor of mouth
Kanno	2020	Previously treated
Lam	2004	T4 Prior excisional biopsy
Mair	2021	NA
Noorlag	2020	NA
Park	2011	Treated for the same diagnosis before hospitalization
Ravikanth	2020	NA
Takamura	2022	T3–T4 Neoadjuvant chemotherapy or radiation therapy
Tang	2022	Other head and neck tumors previously Any prior treatment (biopsy, surgery, radiotherapy, or chemotherapy) for tongue cancer Poor MRI quality due to various factors (movement or artificial implants) MRI done > 2 weeks before surgery
Terada	2020	DOI < 2 mm on pathology

Verma	2019	Previous head-and-neck cancer Prior surgery or radiotherapy to the neck
Yesuratnam	2014	NA

Supplementary file 3: Exclusion criteria of included US studies and US probes used

Author	Year	Exclusion criteria	US probe
Chammas	2011	NA	B-mode intraoral scanning, Gray- scale US system (Logiq 500; GE Medical Systems, Milwaukee, WI) T-shaped transducer (H40212LM/T739 T-Type; GE Medical Systems) 5- to 10-MHz pulsed ultrasonic beam (field-of-view 39 mm) and a 44 9 10-mm linear array
Harada	2021	Treated in another center Peroperative neoadjuvant treatment Very exophytic (>3mm) or remarkably ulcerated (>3mm) tumor	Ultrasound Unit (HI VISION Avius, Hitachi Healthcare Systems, Japan) with 13-MHz hockey-stick type micro-linear probe
Iida	2018	Prior head and neck chemotherapy or radiotherapy Prior glossectomy	16-MHz scanner. T-shaped ultrasonographic probe (model UST-713T/Intraoperative Electronic Linear Probe; Hita- chi Aloka Medical, Ltd., Tokyo, Japan). The aperture of the transducer had a length of 82 mm, width of 15 mm, and depth of 13.5 mm.
Kaneoya	2009	Previously treated	12-MHz linear array transducer (PLM-1202S; Toshiba Medical Systems Co, Tochigi, Japan)
Konishi	2021	Grade not available in pathology report N+ Synchronous cancer T3,T4a	ProSound Alpha 7 surgical ultrasound system (Hitachi Aloka Medical, Tokyo, Japan). 38- or 50-mm linear probe (7.5 MHz)
Kurokawa	2005	NA	Intracavity transducer of 7.5 MHz (Echo Camera SSD-1200CV; Aloka, Tokyo, Japan)
Natori	2008	Lesion not detectable by US Previously treated	B-mode intra-oral US. Gray scale imaging system SSD1200CV (ALOKA Co., LTD.,Tokyo, Japan). 7.5 MHz I-shaped or T-shaped linear transducers
Noorlag	2020	NA	EpiQ 5, with CL15-7 transducer, Philips Medical Systems, Best, The Netherlands
Shintani	1997	NA	Intracavity transducers (PEF-704LA, Toshiba, Tokyo, Japan) 7 MHz ultrasonic beam and a 34 mm linear array
Shintani	2001	NA	7.5-MHz intracavitary transducers (PEF-704LA, Toshiba, UST-995, UST-5536, Aloka, Tokyo, Japan)
Takamura	2022	T3–T4 Neoadjuvant chemotherapy or radiation therapy	HI VISION Preirus (Hitachi, Tokyo, Japan) Hockey stick-type small transducer (EUP-O54J) with a frequency of 7 to 13 MHz
Yamane	2007	Prior chemotherapy or radiotherapy N+	Aloka SSD-630 ultrasound system (Aloka, Tokyo,Japan). 10-MHz mechanical sector transducer (Aloka modified ASR-32 WU-10)
Yesuratnam	2014	NA	15–7 MHz L15-7io linear ultrasound transducer 'hockey stick probe'on a Philips iU22 machine (Philips Medical,Netherlands) with a 23-mm rectangular field of view and single near-field focal zone.

Supplementary file 4: Exclusion criteria of included CT studies

Author	Year	Exclusion criteria
Baba	2021	Dorsal, tip and ventral tongue lesions
Chin	2021	Recurrent cancer Incomplete pathology report Positive margins CT Slice thickness > 3mm CT done > 1 month before surgery
Locatello	2020	Recurrent SCC cT4 with frank bone invasion > 8 weeks delay between imaging and surgery Dorsal tongue lesions correlation excluded; not significative. Only lateral tongue lesions included
Takamura	2022	Neoadjuvant chemotherapy or radiation therapy

Article 2 - Can punch biopsy accurately measure depth of invasion in early oral tongue squamous cell carcinoma? A prospective study.

Authors: Béatrice Voizard MD FRCSC[1,2], Olguta-Ecaterin Gologan MD FRCPC[1,3], Tareck Ayad MD FRCSC[1,2,3], Eric Bissada MD DMD FRCSC[1,2,3], Louis Guertin MD FRCSC[1,2,3], Jean-Claude Tabet MD FRCSC[1,2,3], Guillaume B. Cardin MSc[3], Laurent Létourneau-Guillon MD MSc FRCPC[1,3], Apostolos Christopoulos MD MSc FRCSC[1,2,3]

Institutions:

[1]Faculty of Medicine, Université de Montréal, Montreal, Canada

[2]Division of Otolaryngology – Head and Neck Surgery, Department of Surgery, Université de Montréal, Montreal, Canada

[3]Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, Canada

Corresponding author:

Apostolos Christopoulos, MD MSc FRCSC

a.christopoulos@umontreal.ca

Associate Professor, Division of Otolaryngology-Head and Neck Surgery

Université de Montréal

Permanent address : Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM)

900 St-Denis St

Montreal- QC, Canada, H2X 0A9

Tel: (514) 890-8000 - Fax: (514) 412-7328

Abstract

Objective: The inclusion of depth of invasion (DOI) in the American Joint Committee on Cancer's staging system for oral tongue squamous cell carcinoma (OTSCC) has major clinical impacts. Only few studies have evaluated the accuracy of preoperative biopsy to predict DOI. The primary objective of this prospective study was to evaluate the reliability of preoperative punch biopsy for measurement of DOI in early OTSCC, and to compare it to clinical evaluation by digital palpation. The secondary objective was to evaluate punch biopsy's ability to differentiate between carcinoma in situ (Tis) and invasive carcinoma.

Methods: A deep punch biopsy was used to sample tumors preoperatively in the deepest part of the tumor in patients with early (Tis, T1 or T2, N0) OTSCC. Spearman's correlations were calculated between DOI measured on digital palpation (cDOI), biopsy (bDOI) and final histopathology (pDOI). The sensitivity and specificity of punch biopsy for distinguishing Tis from invasive carcinoma were also calculated.

Results: A total of 27 patients were enrolled. The correlation coefficients between bDOI and pDOI, and between cDOI and pDOI were 0.603 ($CI_{95\%}$=[0.202 – 0.884]) and 0.894 ($CI_{95\%}$=[0.749 – 0.955]), respectively. The sensitivity and specificity of punch biopsy for distinguishing in situ from invasive carcinoma among included patients were 0.89 ($CI_{95\%}$=[0.65-0.99]) and 0.86 ($CI_{95\%}$=[0.42-1.00]), respectively.

Conclusion: Although the correlation between bDOI and pDOI was lower than between cDOI and pDOI, punch biopsy appears to be a reliable tool to distinguish Tis from invasive carcinoma. Larger studies are needed before recommending its use to evaluate pDOI preoperatively or to determine which patients with early OTSCC would benefit from elective neck dissection.

Highlights

Depth of invasion is a prognostic factor in oral tongue squamous cell carcinoma

Depth of invasion measured on punch biopsy is not highly reliable

In carcinoma in situ, punch biopsy is reliable to exclude invasion

Keywords

Head and Neck Cancer, Oral cancer, Tongue Neoplasms, Squamous Cell Carcinoma, Biopsy, Diagnostic Imaging, Neoplasm Staging, Neoplasm Invasiveness

Introduction

In contemporary literature, pathologic DOI (pDOI) has proven to be an important prognostic factor in early oral cavity squamous cell carcinoma (OTSCC). pDOI is defined as the distance from the level of the epithelial basement membrane (BM) to the tumor invasive front[13]. This measure is distinct from tumor thickness (TT), which is defined as the total thickness of the tumor, combining both the invasive depth and the exophytic portion above the BM.[22] Studies show that pDOI correlates with disease-specific survival, as well as occult lymph node metastasis.[13, 74, 75] This has led to its incorporation in the 8th edition of the American Joint Committee on Cancer tumor TNM staging system (AJCC8) (Table 1). Although the optimal cut-off value above which an elective node dissection (END) is indicated has been debated in the literature[14-16, 76], the most recent clinical guidelines from the National Comprehensive Cancer Network (NCCN) now mandate an END for any T1-2, N0 oral cavity SCC with a pDOI above 4mm. Even for pDOI above 2 mm, END should be considered [18]. The NCCN guidelines also state that pDOI should be determined by palpation. Thus far, no study has evaluated the reliability of digital palpation to estimate pDOI in the preoperative setting.

The adequate determination of pDOI in the preoperative setting is critical to proper surgical management. Resection margins depend on pDOI, as they differ between a carcinoma in situ (Tis) (pDOI = 0) and an invasive carcinoma. Moreover, accurate preoperative determination of pDOI can make the difference between performing concomitant glossectomy with END instead of waiting for pDOI to be measured on a gross tumor specimen before performing END in a second surgery. Many recent studies have correlated DOI measured on ultrasound (US)[25, 26], magnetic

resonance imaging (MRI)[77] or computed tomography (CT) [53, 58, 61, 64], with pDOI. Only two studies have evaluated the reliability of DOI measured on biopsy (bDOI) as a predictor of pDOI [19, 20]. The primary objective of this prospective study was to evaluate the reliability of preoperative punch biopsy to measure preoperative pDOI in early (Tis-T1-T2, N0) OTSCC and to compare it to digital palpation. The secondary objective was to evaluate punch biopsy's reliability to distinguish Tis from invasive carcinoma.

Methods

This unicentric prospective diagnostic test accuracy study was conducted between March 2019 and March 2022 at the Centre Hospitalier de l'Université de Montréal (CHUM), a tertiary reference center in Montreal, Canada. This study was approved by the CHUM's Institutional Review Board. Patients were recruited either at the time of diagnosis, or before surgery. All patients were adults (18 years or above) with a probable diagnosis of Tis-T1-T2, N0 OTSCC that had not been previously operated. Other cancers were excluded. Clinical DOI (cDOI) was assessed on palpation by fellowship-trained head & neck surgeons (≥10years in practice). bDOI was measured on a 3 mm or 4 mm diameter punch biopsy specimen. The punches used were standard Integra Miltex biopsy punches that have a maximal biopsy depth of 12mm (7 +/- 5mm) (Integra LifeSciences, Princeton, USA).

For ulcerated or flat lesions, a biopsy was taken in the center of the lesion, where DOI was assumed to be maximal (Figure 1). Of note, the pathologist (author OEG) never had difficulty finding the level of the basement membrane even if the peripheric normal tissue was not included in the biopsy sample. For mildly exophytic lesions (Figure 1.b), biopsy was taken at the site of maximal thickness. There weren't any highly exophytic lesions in the included population. All biopsies were taken at a 90-degree angle with the tumor surface. Patients recruited before ever being biopsied were biopsied in the clinic under local anesthesia. Patients previously

biopsied in other centers were biopsied again at the time of glossectomy, under general anesthesia. pDOI retrieved from the final histopathology reports was used as the gold-standard.

Some biopsy specimens were initially inked at the bottom surface, but after consulting with a subspecialized head and neck pathologist (author OEG), inking was deemed unecessary. Histopathologic analysis of the punch biopsy specimens was performed after formalin fixation, standard paraffin embedding, 4-micron sectioning and hematotoxin and eosin staining. pDOI was measured by drawing a reference line at the basement membrane and drawing a "plumb line" perpendicular to the reference line, down to the tumor-invasive front into the stroma. The same method was used on both biopsy specimens and final glossectomy specimens. (Figure 2)

Analyses were carried out using the R (R Core Team 2019) package meta [46]. DOI was calculated in millimeters. Spearman's correlation coefficients were calculated to compare bDOI to pDOI, cDOI to pDOI and cDOI to bDOI. Missing values were omitted from statistical analysis.

Results

Demographics

Twenty-seven patients were included in the study. Relevant demographic data and tumor specifications are summarized in Table 2. The majority of participants were male (67%). Patients' age ranged from 25 to 97 years old at the time of biopsy (mean = 67.2). Most of the biopsies were done in the operating room (OR) at the time of glossectomy, while only 7 were done in the clinic. For patients whose biopsy was done in clinic (most of which were T1 or Tis), the median delay between bDOI and pDOI measurements was 43 days. The vast majority of lesions (82%) were located on the lateral surface of the tongue, whereas the ventral and dorsal surfaces comprised of only 11% and 7%, respectively. The average greatest diameter of the included

lesions was 16.8 mm. Overall, there were 7 Tis, 10 T1 lesions, 8 T2 lesions and 1 T3 OTSCC. Perineural invasion (PNI) and lymphovascular invasion (LVI) was identified in 8 and 2 specimens, respectively. Biopsy was well tolerated by all enrolled patients. No significant bleeding occurred.

Correlations between DOI measures

Table 3 presents the Spearman's correlation coefficients between cDOI and bDOI, bDOI and pDOI and between cDOI and pDOI. Scatterplots are shown in Figure 3. It appears that the lowest correlation coefficients were between cDOI vs bDOI, and bDOI vs pDOI, with values of 0.743 ($CI_{95\%}$= [0.484 – 0.887]) and 0.603 ($CI_{95\%}$= [0.202 – 0.884]), respectively. In comparison, the correlation coefficient between cDOI and pDOI was higher, at 0.894 ($CI_{95\%}$=[0.749 – 0.955]).

Distinguishing Tis from invasive carcinoma

Table 4 presents the sensitivity, specificity, and predictive values of punch biopsy. As expected, the positive predictive value (PPV) of punch biopsy was high, at 0.94 ($CI_{95\%}$=[0.71-1.00]). In the present study, the sensitivity of punch biopsy was 0.89 ($CI_{95\%}$=[0.65-0.99]). Specificity was also high at 0.86 ($CI_{95\%}$=[0.42-1.00]).

Discussion

Reliability of bDOI

To our knowledge, this is the first prospective proof of concept on punch biopsy's reliability to predict pDOI in early OTSCC. In light of our results, it appears that cDOI correlates better with pDOI than does bDOI.

In the current literature, few studies have evaluated the reliability of bDOI, and most did not use a standardised punch biopsy technique. The two most important considerations when measuring bDOI are to take the sample in a representative area of the tumor, and to include the tumor-invasive front in the biopsy sample. Dhanda *et al*. suggested that a biopsy depth of 10 mm was required to include the tumor invasive front in at least 79% of tumors. Their retrospective review of 139 biopsy samples from various oral cavity subsites showed that only 28% of punch biopsies had a biopsy depth greater than tumor depth[78]. This may be partially attributed to anatomical constraints in other oral cavity subsites like the floor of mouth and alveolar crests. In another retrospective study by Moore *et al*., bDOI could not be measured in 16 of 36 (44%) included biopsy specimens as tumor extend to the full depth of the specimen.[20] The biopsy technique was not described and presumably not standardised. In our study, all samples were deep enough to measure bDOI, as they included the deep tumor-stroma junction. This was the result of the standardised technique used and the depth of the selected punches. Therefore, the inclusion of the tumor-invasive front could not have affected the results and cannot explain the relatively low correlation between bDOI and pDOI.

Potential injuries to Warthin's duct, the floor of mouth structures and airway compromise may be concerns when using a deep punch[78]. Our study focused only on oral tongue tumors where deep punches are generally safe. Although this was not a primary objective of the present study, and the number of included patients is low, punch biopsy in clinic was not harmful, even when taken at a maximal punch depth of 12mm.

Sampling the most invasive part of the lesion may also present a challenge. In our samples, the mean greater diameter of included tumors on final histopathology was 16.8 mm. A 3- or 4-mm punch diameter represents 18% to 24% of this tumor diameter. Therefore, the site of maximal DOI may not be adequately sampled by such a small punch (See Figure 2b). Overall, sampling of the tumor invasive front outside of the site of maximal tumor depth may represent a major

confounder. It may limit punch biopsy's representativeness of pDOI and explain the low correlation.

Another potential factor that can explain our results was the delay between bDOI and pDOI measurements for patients whose biopsy was not taken in the OR. Tumor growth during that time confounds the correlations. Taking biopsies in clinic may reduce the reliability of bDOI, but it maximizes the external validity and clinical applicability of our findings.

In the clinic, infiltration with xylocaine and epinephrine may also have distorted the tumor anatomy and artificially increased bDOI. Xylocaine diffusion within layers of the dermis and hypodermis may increase their height and therefore augment the thickness of the tissue, increasing DOI. Furthermore, tumor shrinkage is a well-studied phenomenon that may arise during the process of formalin fixation and paraffin imbedding. Studies suggest that shrinkage can reduce tumor dimensions between 10% and 24%.[79, 80] The impacts of xylocaine infiltration combined with tumor shrinkage have not been evaluated in the present study. Of note, cDOI was estimated by clinicians before xylocaine infiltration, which could explain why the correlation between cDOI and pDOI was greater than that of bDOI and pDOI. Because the included population was small, no significant subgroup analysis has been performed on patients who were not infiltrated before the punch biopsy.

Another potential confounder was that some biopsies were taken in tumors that had already been biopsied, where scarring and fibrosis might thicken tumoral tissue. Again, these tumors were not excluded in an effort to maximise the external validity of this study, as many tumors arise in areas of previous premalignant changes that may have been biopsied multiple times.

In recent years, systematic reviews and meta-analyses (SRMA) have been published on radiologic measurement of DOI (rDOI) or tumor thickness (rTT). US has mostly been used to measure rTT. In their recent SRMA on the use of US to measure rDOI or rTT, Klein Nulent *et al*. found a pooled Pearson correlation coefficient of 0.88 ($p<0.001$) between US measures and histopatology measures.[81] In another SRMA, Tarabichi *et al*. found a higher pooled correlation coefficient of 0.95 ($p<0.0001$)[82]. As for MRI, an SRMA by Li *et al*. obtained an overall intraclass coefficient of 0.89 ($CI_{95\%}$= [0.837-0.895]) between rTT or rDOI and pTT or pDOI.[77] Overall, these correlation coefficients are considerably higher than the ones we obtained with punch biopsy (bDOI vs pDOI coefficient of 0.603 ($CI_{95\%}$=[[0.202 -0.884])). Interestingly, they do compare with clinical palpation of the tumor, as reflected by our 0.894 ($CI_{95\%}$=[0.749-0.955]) coefficient between cDOI and pDOI.

Distinguishing Tis from invasive carcinoma

The second aim of this study was to evaluate the reliability of punch biopsy to distinguish Tis from invasive carcinoma. To our knowledge, this subject has not been studied prospectively in the literature. Central sampling of the tumor with a deep punch biopsy may be a reliable tool to rule out the presence of invasion and prevent surgeons from resecting larger margins than necessary when proceeding with excision. Maximizing the sensitivity of punch biopsy is of the utmost importance to adequately rule out the presence of invasive carcinoma when the biopsy shows only Tis. This allows the surgeon to not only take narrow excision margins but also to potentially vaporize the lesion instead of proceeding with wide local excision. In the present study, the sensitivity of the punch biopsy was 0.89 ($CI_{95\%}$=[0.65-0.99]). This suggests that a central, deep punch biopsy may be a good tool to rule out the presence of invasion, but a larger sample size would be necessary to reduce the confidence interval. Of note, the punch biopsy in one patient suggested invasive carcinoma, while the final specimen showed only Tis. It is possible that the only area of invasive carcinoma was excised by the biopsy. This explains the PPV lesser than 100% in the present study.

We still advise that any lesion showing significant macroscopic changes, induration, ulceration, rapid growth, or any other clinical signs of invasion should be excised rather than vaporised. Final histopathologic analysis remains the gold standard to distinguish Tis from invasive lesions. In the future, larger studies could ascertain the role of punch biopsy to adequately rule out neoplasm invasiveness.

Finally, some studies that measured rDOI and rTT have previously focused on correlating radiologic measurements not with pDOI or pTT, but directly with other prognostic factors[44, 71]. Few studies on oral cavity punch biopsy's reliability to predict nodal metastasis have been published. Seki *et al.* correlated a bTT of more than 3mm with lymph node metastases (OR=47.5, $CI_{95\%}$=[3.9-511.2]), making it an interesting prognostic tool, irrespective of the correlation between bTT and pTT.[83] Since then, DOI has surpassed TT as a prognostic factor.[13, 75] A direct correlation between bDOI and the number of positive lymph nodes on END would be the most clinically useful tool to implement the use of bDOI in clinical practice. In the present study, only 10 patients had an END, of which only two had a single positive lymph node on final histopathology. These patients had a bDOI of 3- and 4-mm and a pDOI of 5- and 7- mm, respectively. No significant subgroup analysis can be made yet to correlate bDOI to clinical prognostic factors.

Conclusion

To our kowledge, this study is the first prospective study on punch biopsy's reliability to measure pDOI in early OTSCC. It is the first proof-of-concept using standardised deep central punches to measure bDOI. Even though a standardised biopsy technique allowed a deep enough tumor-sampling, bDOI did not correlate strongly with pDOI. Clinical palpation by experienced surgeons correlated better with pDOI and was comparable to the limited data on radiologic measurement of DOI. In the present study, xylocaine infiltration, tissue manipulation, delay between biopsy and final pDOI measurement and sampling outside the deepest tumor site may explain the

results. Interestingly, punch biopsy's sensitivity to rule-out the presence of invasiveness in Tis was 0.89. Subsequent studies on this subject should directly correlate bDOI with clinical outcome measures, as this could prove whether the use of punch biopsy could effectively alter preoperative treatment planning. Larger studies are needed before recommending its use to decide whether elective neck dissection is necessary in early OTSCC and whether it can reliably rule out focal invasiveness in Tis.

Declarations

Availability of data and material

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the Dr Azar-Angelil Research Chair in Head and Neck Oncology of the Université de Montréal (AC).

Author's contribution

Béatrice Voizard MD FRCSC : study design, data collection, manuscript redaction, manuscript revision.

Olguta-Ecaterin Gologan MD FRCPC: data collection, manuscript redaction, manuscript revision

Tareck Ayad MD FRCSC: data collection, manuscript revision

Eric Bissada MD DMD FRCSC: data collection, manuscript revision

Louis Guertin MD FRCSC: data collection, manuscript revision

Jean-Claude Tabet MD FRCSC: data collection, manuscript revision

Guillaume B. Cardin MSc: data collection, manuscript redaction, manuscript revision

Laurent Létourneau-Guillon MD MSc FRCPC : statistical analysis, manuscript redaction, manuscript revision

Apostolos Christopoulos MD MSc FRCSC : study design, data collection, manuscript revision

Table 1. – The American Joint Committee on Cancer Tumor Node Metastasis staging system, eighth edition[5]

T Category	Description
T1	Tumor ≤ 2cm, DOI ≤ 5mm
T2	Tumor ≤ 2cm, DOI > 5mm and ≤10mm or Tumor > 2cm but ≤ 4 cm and DOI ≤ 10mm
T3	Tumor > 2cm but ≤ 4 cm and DOI > 10mm or Tumor > 4 cm and DOI ≤ 10mm
T4a	Tumor > 4 cm and DOI > 10 mm or Tumor invades adjacent structures only
T4b	Tumor invades masticator space, pterygoid plates, or skull base and/or encases the internal carotid artery

Table 2. – Patient demographics and tumor characteristics

Patient population	n = 27
Demographic data	
Age (y, mean ± SD)	67.2 ± 16.0
Male gender	18 (67%)
Location where biopsy performed	
Clinic	7 (26%)
Operating room	20 (74%)
Lesion location	
Dorsum	3 (11%)
Lateral	22 (81%)
Ventral	2 (8%)
Perineural invasion	
Yes	8 (30%)
No	19 (70%)
Lymphovascular invasion	
Yes	2 (8%)
No	25 (92%)
Nodal disease	
No neck END performed	15 (56%)
No positive LN	8 (30%)
Positive LN	2 (8%)

LN : lymph node; LVI : lymphovascular invasion; END : elective neck dissection; PNI : perineural invasion; SD : standard deviation

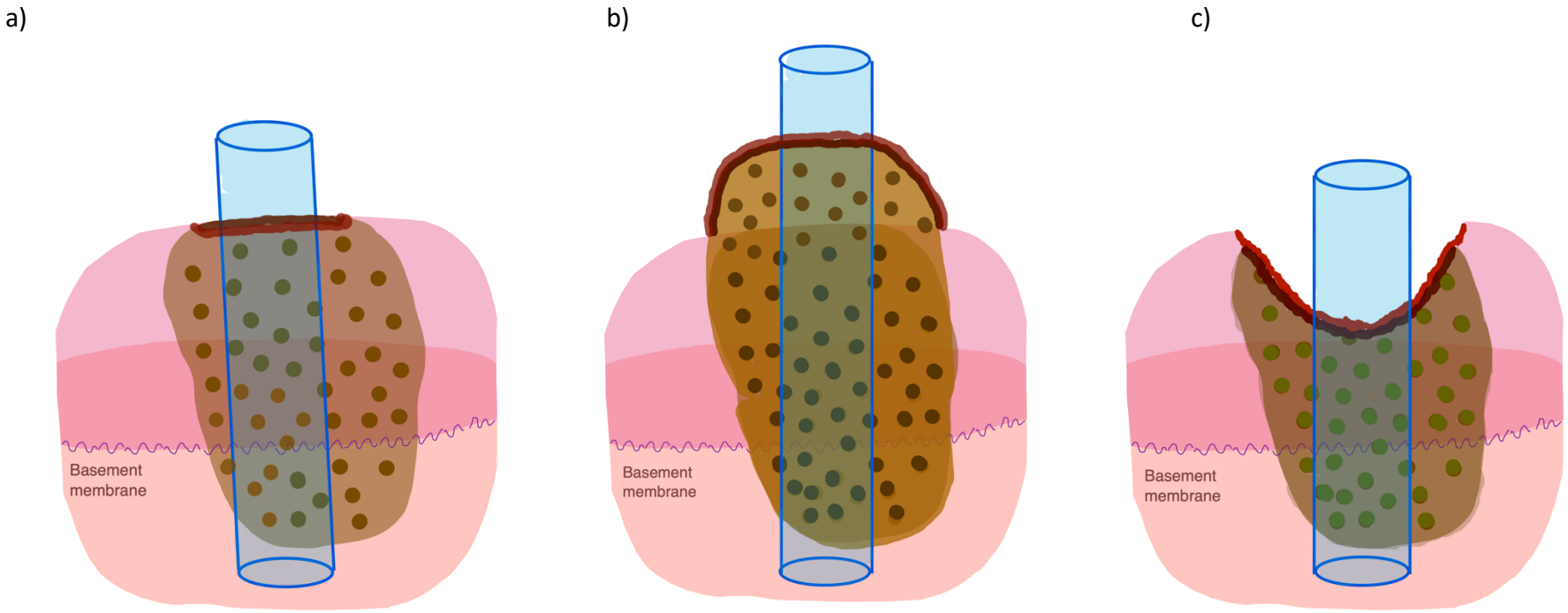
Table 3. – Spearman's correlation coefficients between cDOI, bDOI and pDOI measurements

Comparison	Number of samples	Spearman's correlation coefficient	$CI_{95\%}$
cDOI vs bDOI	22	0.743	[0.484-0.887]
bDOI vs pDOI	25	0.603	[0.202 -0.884]
cDOI vs pDOI	25	0.894	[0.749-0.955]

Table 4. – Accuracy and reliability of punch biopsy to distinguish Tis from invasive carcinoma in suspected Tis-T1-T2, N0 OTSCC (n=25)

	Value	$CI_{95\%}$
Sensitivity	0.89	[0.65-0.99]
Specificity	0.86	[0.42, 1.00]
Positive predictive value	0.94	[0.71, 1.00]
Negative predictive value	0.75	[0.35, 0.97]

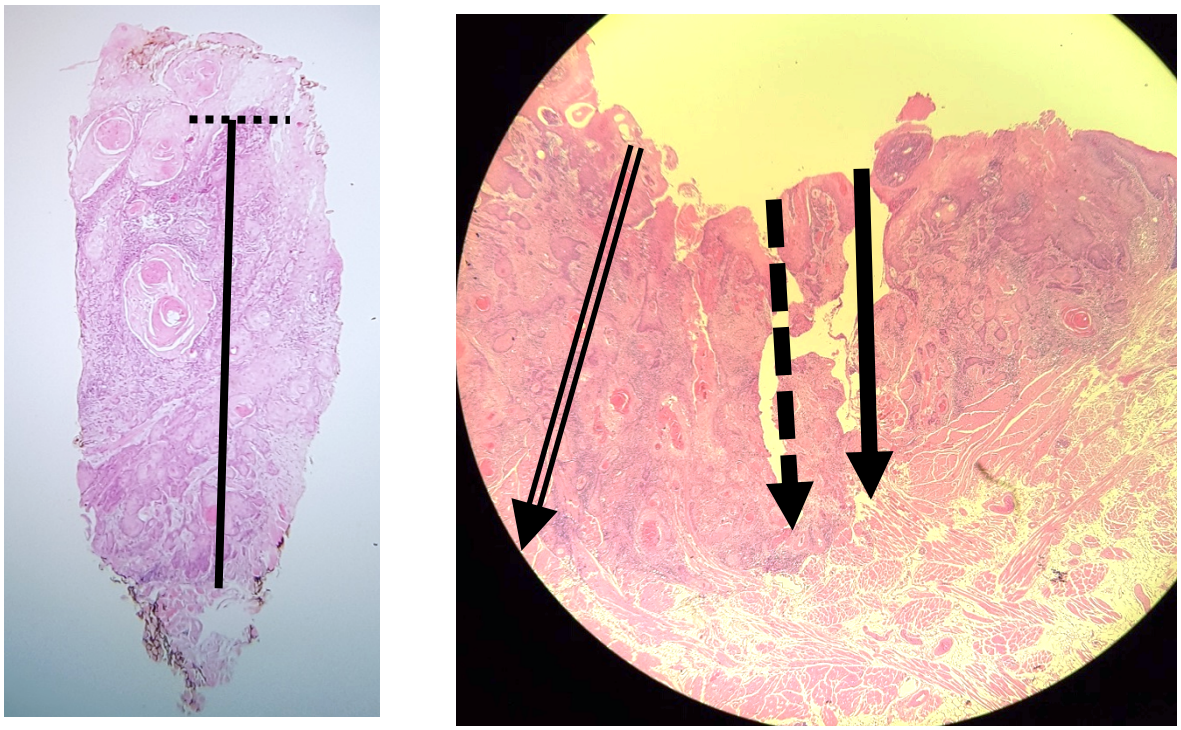
Figure 1. – Punch biopsy technique



Caption: Punch biopsies were taken in the thickest part of the tumor. This was generally the central portion of the tumor, whether it was a) plane, (b) exophytic or (c) ulcerative lesion. Care was taken to insert the punch to its full depth in the tongue to include the tumor invasive front. In highly exophytic lesions, sampling the tumor at the junction between the tumor stalk and the normal tongue surface may optimise DOI measurement, but no highly exophytic lesions were included in this study's patient population.

Figure 2. – pDOI measurement technique.

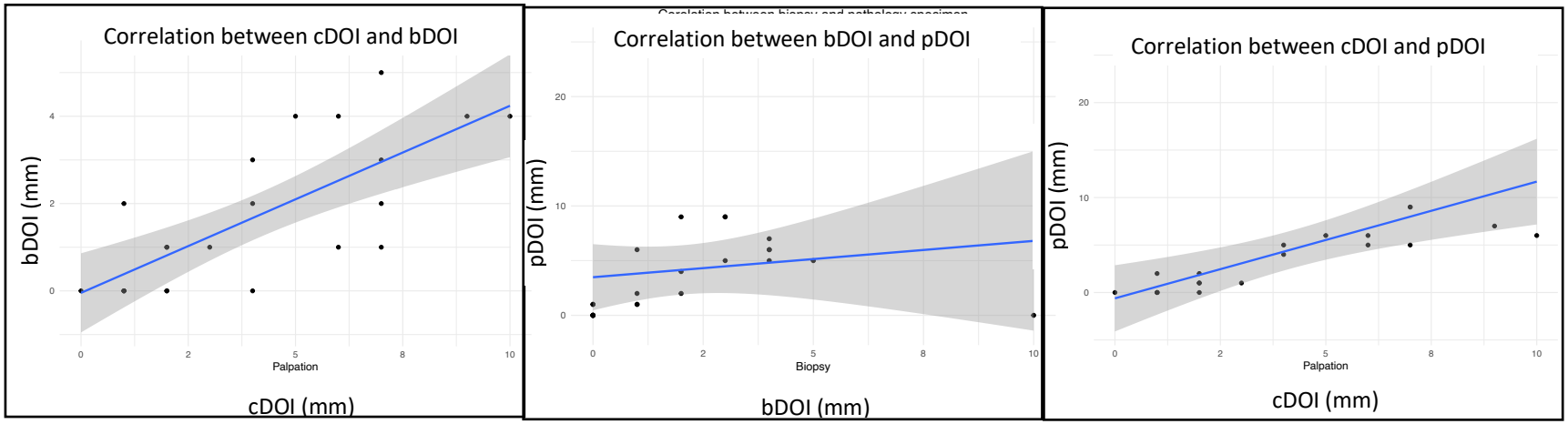
a) b)



Caption : (a) A line is drawn at the junction between the basement membrane and the epithelium (dotted line), and a plumb line (solid line) is then drawn perpendicular to the basement membrane to measure pDOI. (b) The tumor specimen shows the defect created by two punch biopsies taken at 90 angles to the tumor surface. The first punch was taken at partial depth (dash arrow), while the second one was taken at full depth (solid arrow). This picture illustrates that sampling with a punch biopsy may not be done at the deepest point of the tumor. Here, the point of maximal tumor depth was not sampled and is located at double line arrow.

Figure 3. – Scatterplots showing the correlations between cDOI, bDOI and pDOI

a) b) c)



Caption: Scatterplots showing correlations between (a) cDOI and bDOI (b) bDOI and pDOI and (c) cDOI and pDOI. cDOI and pDOI show a stronger correlation

Chapitre 4 – Discussion générale

Tel que discuté dans les deux articles, l'évaluation préopératoire de la PI des carcinomes épidermoïdes de la langue mobile est limitée par de multiples facteurs. Puisque les modalités d'imageries ont été étudiées sur des populations hétérogènes et que les protocoles d'imagerie ne sont pas standardisés, l'implantation de ces méthodes dans la pratique clinique présente de multiples défis. Par ailleurs, la biopsie au poinçon semble fournir une mesure de PI qui ne corrèle pas fortement avec la mesure histopathologique finale, mais elle semble être un outil sensible pour distinguer le Tis de l'invasif. Finalement, la majorité des études corrélant les PI radiologiques ou biopsiques aux valeurs de PI histopathologiques n'établissent pas de corrélation directe avec les facteurs de mauvais pronostic tels que la présence de métastases nodales occultes ou encore l'invasion lymphovasculaire ou périneurale. Ceci diminue l'applicabilité clinique des résultats de ces études.

De plus, bien que la PI corrèle avec le pronostic et fasse maintenant partie du AJCC8, elle n'est pas pour autant un paramètre infaillible. Par exemple, Berdugo *et al.* ont souligné l'importance de considérer les foyers extra-tumoraux de carcinome comme des facteurs importants dans l'évaluation histopathologique. Leur impact sur la méthode de mesure de la PI n'est pas bien définie dans AJCC8, ni dans les études contemporaines. En outre, ils soulignent que la résection du foyer d'invasion le plus profond lors de la biopsie peut empêcher la mesure finale de PI sur le spécimen pathologique.[84] Tel que mentionné dans quelques articles, le phénomène de rétrécissement des spécimens réséqués lors des manipulations en pathologie (*tumor shrinkage*) pose un autre problème, spécialement lorsqu'une différence de seulement quelques millimètres dans la mesure obtenue peut avoir un impact significatif sur l'indication d'évidement cervical électif.[54] À la lumière de ces limites, plusieurs chercheurs ont récemment incités la communauté scientifique à se référer à d'autres mesures que la PI seule pour pronostiquer les tumeurs.

Parmi ceux-ci, on note le *tumor budding*; la présence de cellules tumorales isolées ou en amas dans le stroma sain en aval du front invasif. Seki *et al.* ont établi une corrélation forte entre

le *tumor budding* évalué sur le spécimen de biopsie et la survie à long terme. [85] D'autres études ont également évalué l'impact pronostic du grade histopathologique au site où la PI est maximale. [86] On en conclue que, si la valeur pronostique de la PI est déjà démontrée depuis longtemps, celle-ci pourrait être supplémentée d'autres paramètres histopathologiques dans le futur pour mieux classifier les tumeurs. La mesure de ces paramètres en préopératoire s'accompagnerait potentiellement d'autres défis.

Il importe de rappeler que les deux articles ainsi que la majorité des références précédemment citées se rapportent principalement à la langue mobile. Toutefois, la PI est partie intégrante du *staging* d'AJCC8 pour tous les sous-sites de la cavité orale. L'évaluation de la PI dans ces autres sous-sites demeure moins étudiée et serait un champ intéressant d'études ultérieures. Finalement, la performance clinique de la PI en tant que paramètre pronostic pourra être démontrée par des études populationnelles rétrospectives plusieurs années après l'implantation du AJCC8, qui demeure un système très récent et qui est actuellement basé uniquement sur des études rétrospectives.

Chapitre 5 – Conclusion

Le poids épidémiologique des carcinomes épidermoïdes de la langue mobile est majeur, et l'inclusion de la PI dans l'AJCC8 a donc des impacts cliniques importants. Bien que les valeurs de PI les plus fortement prédictives de mauvais pronostic soient débattues dans la littérature et que d'autres paramètres cliniques jouent également un rôle pronostic clairement établi, la mesure préopératoire de la PI demeure cruciale. Elle permet aux chirurgiens de décider d'emblée de réaliser un évidement cervical électif lorsqu'elle est supérieure à 2 à 4mm[§] et, dans le cas du Tis, permet de minimiser les marges de résection chirurgicale. Ceci peut éviter aux patients de subir deux chirurgies plutôt qu'une et permettre de maximiser leur récupération fonctionnelle post-glossectomie.

Ce mémoire de maîtrise visait à colliger et comparer les connaissances actuelles sur les méthodes de mesure de PI en préopératoire et a conclu que l'IRM semble actuellement l'outil le mieux étudié et le plus fiable pour mesurer la PI. De plus, ce mémoire visait à étudier prospectivement la biopsie au poinçon et a démontré que la PI évaluée sur celle-ci ne semble pas corréler plus fortement avec la PI histopathologique que la palpation clinique. Toutefois, dans les cas de Tis suspectés, la biopsie peut éliminer la présence d'invasion avec une forte sensibilité.

Dans le futur, la standardisation des modalités et protocoles d'imagerie clinique pourrait permettre d'intégrer les mesures radiologiques de la PI à la pratique clinique. D'autre part, l'étude de la biopsie au poinçon sur de plus nombreux échantillons de tumeurs pourrait permettre d'améliorer la puissance statistique des résultats précédemment présentés. Finalement, l'établissement d'une corrélation directe entre les mesures préopératoires de la PI et les facteurs pronostics cliniques, ainsi que l'étude des techniques de mesure de PI sur d'autres sous-sites de la cavité orale pourraient éventuellement maximiser l'impact clinique de nos trouvailles.

[§] Selon les recommandations du NCCN. Différents centres d'expertise en oncologie cervico-faciale peuvent utiliser d'autres valeurs selon les guides de pratique reconnus dans différents pays.

Références bibliographiques

1. *PRISMA 2020.* J Clin Epidemiol, 2021. **134**: p. A5-a6.
2. Patterson, R.H., et al., *Global Burden of Head and Neck Cancer: Economic Consequences, Health, and the Role of Surgery.* Otolaryngol Head Neck Surg, 2020. **162**(3): p. 296-303.
3. Siegel, R.L., et al., *Cancer statistics, 2022.* CA Cancer J Clin, 2022. **72**(1): p. 7-33.
4. *Statistics Canada. Table 13-10-0747-01 Number of new cases and age-standardized rates of primary cancer (based on the November 2017 CCR tabulation file), by cancer type and sex.*
5. Cancer, A.J.C.o., *AJCC Cancer Staging Manual*. Eigth ed. 2017, Chicago, IL: Springer Nature. 1024.
6. Lydiatt, W.M., 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**(2): p. 122-137.
7. Brockhoff, H.C., 2nd, et al., *Correlating the depth of invasion at specific anatomic locations with the risk for regional metastatic disease to lymph nodes in the neck for oral squamous cell carcinoma.* Head Neck, 2017. **39**(5): p. 974-979.
8. Pentenero, M., S. Gandolfo, and M. Carrozzo, *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**(12): p. 1080-91.
9. Chen, W.L., et al., *MRI-derived tumor thickness: an important predictor of outcome for T4a-staged tongue carcinoma.* Eur Arch Otorhinolaryngol, 2012. **269**(3): p. 959-63.
10. Tan, W.J., et al., *Prognostic significance of invasion depth in oral tongue squamous cell carcinoma.* ORL J Otorhinolaryngol Relat Spec, 2012. **74**(5): p. 264-70.
11. Shinn, J.R., et al., *Cumulative incidence of neck recurrence with increasing depth of invasion.* Oral Oncol, 2018. **87**: p. 36-42.
12. Sparano, A., et al., *Multivariate predictors of occult neck metastasis in early oral tongue cancer.* Otolaryngol Head Neck Surg, 2004. **131**(4): p. 472-6.
13. Ebrahimi, A., 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**(12): p. 1138-48.
14. Tam, S., et al., *Depth of invasion as a predictor of nodal disease and survival in patients with oral tongue squamous cell carcinoma.* Head Neck, 2019. **41**(1): p. 177-184.
15. Tarsitano, A., et al., *Tumor Infiltration Depth as Predictor of Nodal Metastasis in Early Tongue Squamous Cell Carcinoma.* J Oral Maxillofac Surg, 2016. **74**(3): p. 523-7.
16. Melchers, L.J., et al., *Tumour infiltration depth >/=4 mm is an indication for an elective neck dissection in pT1cN0 oral squamous cell carcinoma.* Oral Oncol, 2012. **48**(4): p. 337-42.

17. D'Cruz, A.K., et al., *Elective versus Therapeutic Neck Dissection in Node-Negative Oral Cancer.* N Engl J Med, 2015. **373**(6): p. 521-9.
18. Network, N.C.C. *Head and Neck Cancer (Version 1.2022)*. 2022 2022-02-10]; Available from: https://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf.
19. Almangush, A., et al., *Evaluation of the budding and depth of invasion (BD) model in oral tongue cancer biopsies.* Virchows Arch, 2018. **472**(2): p. 231-236.
20. Moor, J.W., et al., *Biopsy examination of squamous cell carcinoma of the tongue: source of significant prognostic information?* Br J Oral Maxillofac Surg, 2010. **48**(8): p. 594-7.
21. Yesuratnam, A., et al., *Preoperative evaluation of oral tongue squamous cell carcinoma with intraoral ultrasound and magnetic resonance imaging-comparison with histopathological tumour thickness and accuracy in guiding patient management.* Int J Oral Maxillofac Surg, 2014. **43**(7): p. 787-94.
22. Moore, C., J.G. Kuhns, and R.A. Greenberg, *Thickness as prognostic aid in upper aerodigestive tract cancer.* Arch Surg, 1986. **121**(12): p. 1410-4.
23. Berdugo, J., et al., *Measuring Depth of Invasion in Early Squamous Cell Carcinoma of the Oral Tongue: Positive Deep Margin, Extratumoral Perineural Invasion, and Other Challenges.* Head Neck Pathol, 2019. **13**(2): p. 154-161.
24. Li, M., Z. Yuan, and Z. Tang, *The accuracy of magnetic resonance imaging to measure the depth of invasion in oral tongue cancer: a systematic review and meta-analysis.* Int J Oral Maxillofac Surg, 2021.
25. Tarabichi, O., et al., *Utility of intraoral ultrasound in managing oral tongue squamous cell carcinoma: Systematic review.* Laryngoscope, 2019. **129**(3): p. 662-670.
26. Klein Nulent, T.J.W., et al., *Intraoral ultrasonography to measure tumor thickness of oral cancer: A systematic review and meta-analysis.* Oral Oncol, 2018. **77**: p. 29-36.
27. Helbig, M., et al., *Intraoperative B-mode endosonography of tongue carcinoma.* Head Neck, 2001. **23**(3): p. 233-7.
28. Lim, Y.C., et al., *Overexpression of c-Met promotes invasion and metastasis of small oral tongue carcinoma.* Oral Oncol, 2012. **48**(11): p. 1114-9.
29. Iwai, H., et al., *Magnetic resonance determination of tumor thickness as predictive factor of cervical metastasis in oral tongue carcinoma.* Laryngoscope, 2002. **112**(3): p. 457-61.
30. Jayasankaran, S.C., et al., *Magnetic resonance imaging: A predictor of pathological tumor dimensions in carcinoma of anterior two-thirds of tongue - A prospective evaluation.* Indian J Cancer, 2017. **54**(3): p. 508-513.
31. Madana, J., et al., *Computerized tomography based tumor-thickness measurement is useful to predict postoperative pathological tumor thickness in oral tongue squamous cell carcinoma.* J Otolaryngol Head Neck Surg, 2015. **44**: p. 49.
32. Mao, M.H., et al., *Accuracy of magnetic resonance imaging in evaluating the depth of invasion of tongue cancer. A prospective cohort study.* Oral Oncol, 2019. **91**: p. 79-84.
33. Mark Taylor, S., et al., *Is preoperative ultrasonography accurate in measuring tumor thickness and predicting the incidence of cervical metastasis in oral cancer?* Oral Oncol, 2010. **46**(1): p. 38-41.
34. Minamitake, A., et al., *Can MRI-derived depth of invasion predict nodal recurrence in oral tongue cancer?* Oral Radiol, 2021. **37**(4): p. 641-646.

35. Moreno, K.F., et al., *Using 3 Tesla magnetic resonance imaging in the pre-operative evaluation of tongue carcinoma.* J Laryngol Otol, 2017. **131**(9): p. 793-800.
36. Murakami, R., et al., *Reliability of MRI-Derived Depth of Invasion of Oral Tongue Cancer.* Acad Radiol, 2019. **26**(7): p. e180-e186.
37. Preda, L., et al., *Relationship between histologic thickness of tongue carcinoma and thickness estimated from preoperative MRI.* Eur Radiol, 2006. **16**(10): p. 2242-8.
38. Ren, J., Y. Yuan, and X. Tao, *Histogram analysis of diffusion-weighted imaging and dynamic contrast-enhanced MRI for predicting occult lymph node metastasis in early-stage oral tongue squamous cell carcinoma.* Eur Radiol, 2021.
39. Saenthaveesuk, P., et al., *Development and validation of multiparametric MRI-based nomogram for predicting occult metastasis risk in early tongue squamous cell carcinoma.* BMC Cancer, 2021. **21**(1): p. 408.
40. Shoukat, S., et al., *Correlation Of Preoperative Volume Of Oral Tongue Squamous Cell Carcinoma (SCC) On CT Scan With Postsurgical Tumour Size.* J Ayub Med Coll Abbottabad, 2021. **33**(3): p. 462-466.
41. Tsushima, N., et al., *The role of prophylactic neck dissection and tumor thickness evaluation for patients with cN0 tongue squamous cell carcinoma.* Eur Arch Otorhinolaryngol, 2016. **273**(11): p. 3987-3992.
42. Vidiri, A., et al., *The role of MRI-derived depth of invasion in staging oral tongue squamous cell carcinoma: inter-reader and radiological-pathological agreement.* Acta Radiol, 2020. **61**(3): p. 344-352.
43. Wang, F., et al., *Magnetic Resonance Imaging-Based Radiomics Features Associated with Depth of Invasion Predicted Lymph Node Metastasis and Prognosis in Tongue Cancer.* J Magn Reson Imaging, 2021.
44. Xu, C., et al., *Significance of depth of invasion determined by MRI in cT1N0 tongue squamous cell carcinoma.* Sci Rep, 2020. **10**(1): p. 4695.
45. Yuen, A.P.W., et al., *Preoperative measurement of tumor thickness of oral tongue carcinoma with intraoral ultrasonography.* Head and Neck, 2008. **30**(2): p. 230-234.
46. (2019), R.C.T. *R: A Language and Environment for Statistical Computing*. 2019 [cited 2019 2019-12-29]; Available from: https://www.R-project.org/.
47. Balduzzi, S., G. Rücker, and G. Schwarzer, *How to perform a meta-analysis with R: a practical tutorial.* Evid Based Ment Health, 2019. **22**(4): p. 153-160.
48. Hedges, L.V.a.O., Ingram, *Statistical Methods for Meta-Analysis*, in *Statistical Methods for Meta-Analysis*. 1985, Academic Press: San Diego.
49. IntHout, J., et al., *Plea for routinely presenting prediction intervals in meta-analysis.* BMJ Open, 2016. **6**(7): p. e010247.
50. Cochran, W.G., *The comparison of percentages in matched samples.* Biometrika, 1950. **37**(3-4): p. 256-66.
51. Higgins, J.P., et al., *Measuring inconsistency in meta-analyses.* Bmj, 2003. **327**(7414): p. 557-60.
52. Viechtbauer, W. and M.W. Cheung, *Outlier and influence diagnostics for meta-analysis.* Res Synth Methods, 2010. **1**(2): p. 112-25.

53. Baba, A., et al., *Usefulness of contrast-enhanced CT in the evaluation of depth of invasion in oral tongue squamous cell carcinoma: comparison with MRI.* Oral Radiol, 2021. **37**(1): p. 86-94.
54. Harada, H., et al., *MRI before biopsy correlates with depth of invasion corrected for shrinkage rate of the histopathological specimen in tongue carcinoma.* Scientific reports, 2021. **11**(1): p. 20992.
55. Noorlag, R., et al., *Assessment of tumour depth in early tongue cancer: Accuracy of MRI and intraoral ultrasound.* Oral Oncol, 2020. **110**: p. 104895.
56. Takamura, M., et al., *A comparative study between CT, MRI, and intraoral US for the evaluation of the depth of invasion in early stage (T1/T2) tongue squamous cell carcinoma.* Oral Radiol, 2022. **38**(1): p. 114-125.
57. Kanno, M., et al., *Comparison of diagnostic accuracy between [(18)F]FDG PET/MRI and contrast-enhanced MRI in T staging for oral tongue cancer.* Ann Nucl Med, 2020. **34**(12): p. 952-959.
58. Chin, S.Y., et al., *Correlation and accuracy of contrast-enhanced computed tomography in assessing depth of invasion of oral tongue carcinoma.* International Journal of Oral and Maxillofacial Surgery, 2021. **50**(6): p. 718-724.
59. Alsaffar, H.A., et al., *Correlation between clinical and MRI assessment of depth of invasion in oral tongue squamous cell carcinoma.* J Otolaryngol Head Neck Surg, 2016. **45**(1): p. 61.
60. Tang, W., et al., *Assessment of tumor depth in oral tongue squamous cell carcinoma with multiparametric MRI: correlation with pathology.* Eur Radiol, 2022. **32**(1): p. 254-261.
61. Locatello, L.G., et al., *A critical evaluation of computed tomography-derived depth of invasion in the preoperative assessment of oral cancer staging.* Oral Oncol, 2020. **107**: p. 104749.
62. Fu, J.Y., et al., *Assessing the magnetic resonance imaging in determining the depth of invasion of tongue cancer.* Oral Diseases, 2021. **27**(3): p. 457-463.
63. Moor, J.W., et al., *Biopsy examination of squamous cell carcinoma of the tongue: Source of significant prognostic information?* British Journal of Oral and Maxillofacial Surgery, 2010. **48**(8): p. 594-597.
64. Takamura, M., et al., *A comparative study between CT, MRI, and intraoral US for the evaluation of the depth of invasion in early stage (T1/T2) tongue squamous cell carcinoma.* Oral Radiology, 2022. **38**(1): p. 114-125.
65. Mair, M., et al., *Diagnostic accuracy of magnetic resonance imaging in detecting depth of invasion of tongue cancers.* Br J Oral Maxillofac Surg, 2021.
66. Koo, T.K. and M.Y. Li, *A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research.* J Chiropr Med, 2016. **15**(2): p. 155-63.
67. Howe, T.E., et al., *Accuracy of staging of oral squamous cell carcinoma of the tongue: should incisional biopsy be done before or after magnetic resonance imaging?* Br J Oral Maxillofac Surg, 2017. **55**(3): p. 298-299.
68. Shintani, S., et al., *Intraoral ultrasonography is useful to evaluate tumor thickness in tongue carcinoma.* Am J Surg, 1997. **173**(4): p. 345-7.
69. Imai, T., et al., *Retrospective observational study of occult cervical lymph-node metastasis in T1N0 tongue cancer.* Japanese Journal of Clinical Oncology, 2017. **47**(2): p. 130-136.

70. Fu, J.Y., et al., *Assessing the magnetic resonance imaging in determining the depth of invasion of tongue cancer.* Oral Dis, 2021. **27**(3): p. 457-463.
71. Natori, T., et al., *Usefulness of intra-oral ultrasonography to predict neck metastasis in patients with tongue carcinoma.* Oral Dis, 2008. **14**(7): p. 591-9.
72. Xu, C., et al., *Significance of depth of invasion determined by MRI in cT1N0 tongue squamous cell carcinoma.* Scientific reports, 2020. **10**(1): p. 4695.
73. Kurokawa, H., et al., *Preoperative ultrasound assessment of tumour thickness in tongue carcinomas.* Asian Journal of Oral and Maxillofacial Surgery, 2005. **17**(3): p. 173-178.
74. Almangush, A., et al., *Depth of invasion, tumor budding, and worst pattern of invasion: prognostic indicators in early-stage oral tongue cancer.* Head Neck, 2014. **36**(6): p. 811-8.
75. Kane, S.V., et al., *Depth of invasion is the most significant histological predictor of subclinical cervical lymph node metastasis in early squamous carcinomas of the oral cavity.* Eur J Surg Oncol, 2006. **32**(7): p. 795-803.
76. Shin, J.H., et al., *Analyzing the factors that influence occult metastasis in oral tongue cancer.* J Korean Assoc Oral Maxillofac Surg, 2020. **46**(2): p. 99-107.
77. Li, M., Z. Yuan, and Z. Tang, *The accuracy of magnetic resonance imaging to measure the depth of invasion in oral tongue cancer: a systematic review and meta-analysis.* International Journal of Oral and Maxillofacial Surgery, 2021.
78. Dhanda, J., et al., *Features and prognostic utility of biopsy in oral squamous cell carcinoma.* Head Neck, 2016. **38 Suppl 1**: p. E1857-62.
79. Pangare, T.B., et al., *Effect of Formalin Fixation on Surgical Margins in Patients With Oral Squamous Cell Carcinoma.* J Oral Maxillofac Surg, 2017. **75**(6): p. 1293-1298.
80. Umstattd, L.A., et al., *Shrinkage in oral squamous cell carcinoma: An analysis of tumor and margin measurements in vivo, post-resection, and post-formalin fixation.* Am J Otolaryngol, 2017. **38**(6): p. 660-662.
81. Klein Nulent, T.J.W., et al., *Intraoral ultrasonography to measure tumor thickness of oral cancer: A systematic review and meta-analysis.* Oral Oncology, 2018. **77**: p. 29-36.
82. Tarabichi, O., et al., *Utility of intraoral ultrasound in managing oral tongue squamous cell carcinoma: Systematic review.* Laryngoscope, 2018.
83. Seki, M., et al., *Histologic assessment of tumor budding in preoperative biopsies to predict nodal metastasis in squamous cell carcinoma of the tongue and floor of the mouth.* Head Neck, 2016. **38 Suppl 1**: p. E1582-90.
84. Berdugo, J., et al., *Measuring Depth of Invasion in Early Squamous Cell Carcinoma of the Oral Tongue: Positive Deep Margin, Extratumoral Perineural Invasion, and Other Challenges.* Head Neck Pathol, 2018.
85. Seki, M., et al., *Tumour budding evaluated in biopsy specimens is a useful predictor of prognosis in patients with cN0 early stage oral squamous cell carcinoma.* Histopathology, 2017. **70**(6): p. 869-879.
86. Kurokawa, H., et al., *The high prognostic value of the histologic grade at the deep invasive front of tongue squamous cell carcinoma.* J Oral Pathol Med, 2005. **34**(6): p. 329-33.