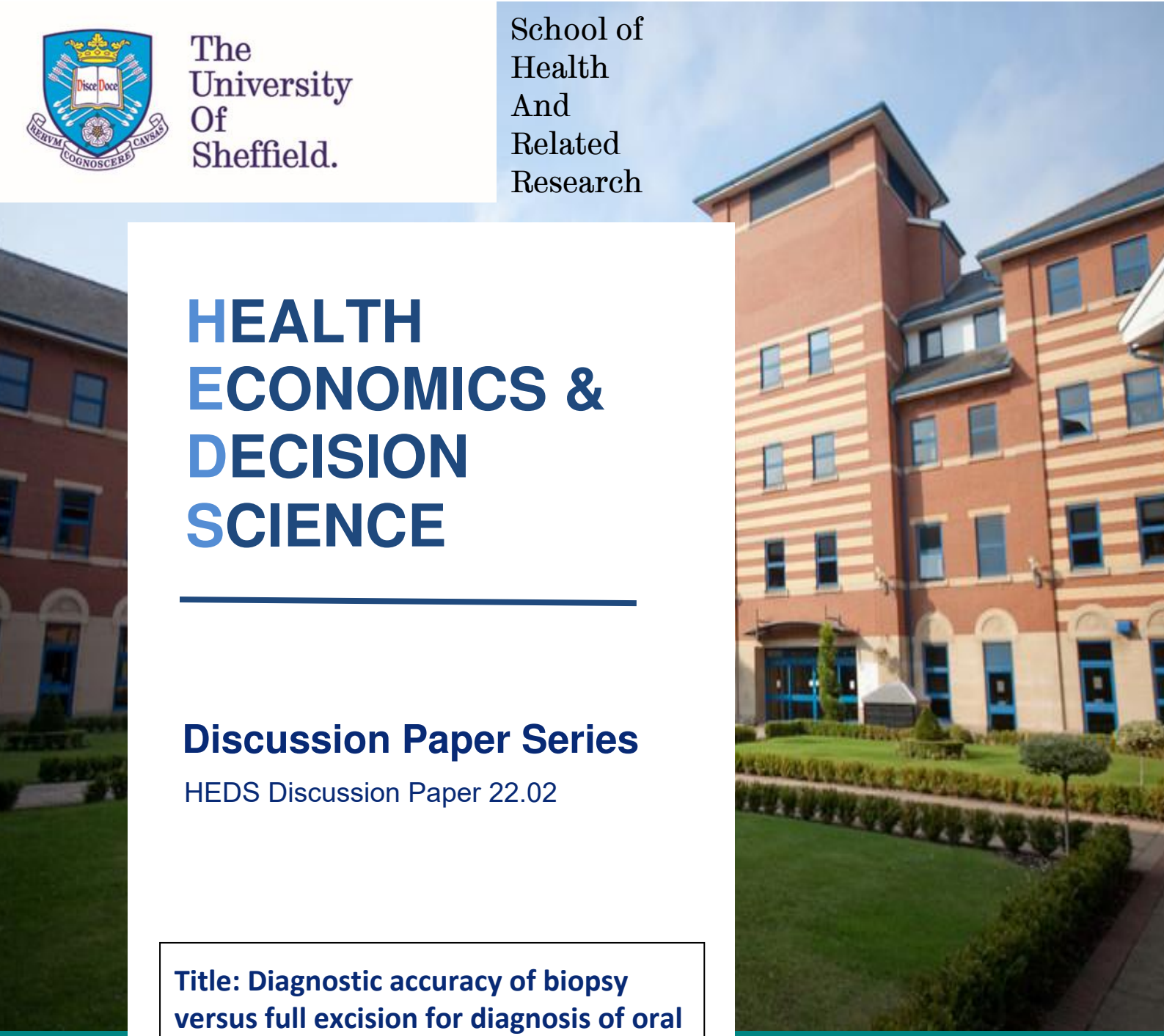
The University Of Sheffield.

School of Health And Related Research



HEALTH ECONOMICS & DECISION SCIENCE

Discussion Paper Series

HEDS Discussion Paper 22.02

Title: Diagnostic accuracy of biopsy versus full excision for diagnosis of oral cancer: A systematic review

Authors:

Katy Cooper Orcid ID:0000-0002-7702-8103
Munira Essat Orcid ID:0000-0003-2397-402X
Alice Bessey Orcid ID:0000-0001-7677-0309
Mark Clowes Orcid ID:0000-0002-5582-9946
Jim Chilcott Orcid ID:0000-0003-1231-7817
Keith D Hunter Orcid ID:0000-0002-7873-0877

Corresponding author: Katy Cooper, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK. Tel: +44 (0) 114 222 0773, Email: k.l.cooper@sheffield.ac.uk

Diagnostic accuracy of biopsy versus full excision for diagnosis of oral cancer: A systematic review

Katy Cooper[a], Munira Essat[a], Alice Bessey[a], Mark Clowes[a], Jim Chilcott[a], Keith D Hunter.[b]

[a]ScHARR, The University of Sheffield, Sheffield, UK.

[b]School of Clinical Dentistry, The University of Sheffield, UK.

Katy Cooper, k.l.cooper@sheffield.ac.uk

Munira Essat, m.essat@sheffield.ac.uk

Alice Bessey, a.r.bessey@sheffield.ac.uk

Mark Clowes, m.clowes@sheffield.ac.uk

Jim Chilcott, j.b.chilcott@sheffield.ac.uk

Keith D Hunter, keith.hunter@liverpool.ac.uk

Corresponding author: Katy Cooper, ScHARR, The University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, telephone +44 (0) 114 222 0773, email: k.l.cooper@sheffield.ac.uk, https://orcid.org/0000-0002-7702-8103

Acknowledgements and funding: We would like to thank the team at Zilico Ltd. for their support throughout the project. This work was supported by funding from the Small Business Research Initiative (SBRI).

Conflicts of interest: The authors have no conflicts of interest.

Abstract

Objectives: To review studies assessing diagnostic accuracy of biopsy of oral lesions for diagnosis of oral cancer and potentially malignant disorders, compared with full excision.

Methods: Systematic review: four databases were searched for studies conducted within developed countries and published between 2009 and January 2020.

Results: Six studies met inclusion criteria, with wide variation in methods and results. For identifying dysplasia or malignancy, sensitivity of biopsy (versus excision) ranged from 42% to 86% and specificity from 75% to 100% across two studies, while concordance varied from 27% to 89% across four studies. For identifying malignant-only lesions, sensitivity was 71% and 94% in two studies, while specificity ranged from 17% to 100% across four studies.

Conclusions: There are few published studies assessing biopsy accuracy, with varying results. Further research should evaluate factors impacting accuracy, such as biopsy depth; multiple biopsies of large lesions; discordance between pathologists; and regular follow-up.

Keywords

Oral cancer, biopsy, diagnostic tests, sensitivity and specificity, systematic review

Background

Oral squamous cell carcinoma (OSCC) is one of the most highly prevalent cancers worldwide, with a worldwide incidence of more than 300,000 cases per year [1]. OSCC has been estimated to have a five year survival rate of around 75% for stage I disease but only around 30% at stage IV of the disease [2, 3]. Once an oral lesion is identified as suspicious through conventional oral examination, the usual management approach is biopsy of the lesion followed by histological examination. However, the diagnostic accuracy of biopsy remains unclear. It is well recognized that the oral mucosa may undergo extensive genetic "field change" which can make the selection of the most appropriate biopsy site very challenging [4]. One possible assessment of biopsy accuracy is comparison with the pathology of full excision of the lesion, if that is available. This allows for assessment of how representative the biopsy pathology is of the whole lesion.

We report a systematic review of the diagnostic accuracy of biopsy in the diagnosis of oral cancer and potentially malignant (dysplastic) disorders, compared with the final pathology of full lesion excision. Our aims were to summarize the available evidence, as well as to highlight the lack of research in this area and to make recommendations for future research.

Methods

A search was conducted in MEDLINE, EMBASE, CINAHL and the Cochrane Library (including the CENTRAL register of controlled trials and Cochrane Database of Systematic Reviews) from inception to January 2020. The search strategy combined terms for oral cancer/dysplasia, terms for biopsy and terms for diagnostic accuracy. Studies were included if they compared accuracy of biopsy versus excision, were conducted in developed countries with similar health systems (Western Europe, North America, Canada, Australia, Japan, South Korea), and were published from 2009 onwards. A second reviewer checked a sample of titles/abstracts, all full texts, and all extracted data. Diagnostic accuracy was assessed for two clinical outcomes: 1) diagnosis of malignant and dysplastic lesions and 2) diagnosis of malignant lesions alone.

Results

Six studies were identified which compared the accuracy of biopsy versus full excision, undertaken in Denmark,[5] Italy,[6] Australia,[7] USA[8], UK [9] and South Korea [10]. Prevalence of dysplastic or malignant lesions varied across the five studies reporting this (26%, 76%, 80%, 88% and 100%), as did prevalence of malignant-only lesions (7%, 61%, 67% and 85%). The time from biopsy to excision was reported in five studies (<1 month, or mean/median of 1.6, 2.5, 4.4 or 10.4 months); longer delays may have allowed lesions to progress between biopsy and excision.

Sensitivity and specificity for dysplastic or malignant lesions

Two studies reported data allowing calculation of sensitivity and specificity of biopsy vs. excision for identification of dysplastic or malignant lesions (**Table 1**). One study was conducted in Denmark and included 101 lesions, with a prevalence of 76% for dysplasia or malignancy on full excision [5]. The other was conducted in South Korea and included 15 lesions, with a prevalence of 80% for dysplasia or malignancy on full excision [10]. The sensitivity of biopsy (compared to excision) for detection of dysplasia or malignancy was 86% in the Danish study and 42% in the Korean study, while the specificity for detection of dysplasia or malignancy was 75% in the Danish study and 100% in the Korean study.

Table 1: Biopsy vs. full excision: Sensitivity and specificity for diagnosis of dysplastic or malignant lesions

Author, Year **Country** **Study Design**	**Study population**	**Mean age (yr)**	**Male (%)**	**Time: biopsy to excision (mean)**	**Definition of positive/ negative case** (via reference standard)	**Prevalence: dysplasia or malignancy**	**TP**	**FN**	**FP**	**TN**	**Sens**	**Spec**	**PPV**	**NPV**
Holmstrup 2007 [5] Denmark Retrospective	Oral mucosal lesions (where biopsy showed epithelial dysplasia and/or lesions on lateral/ventral tongue or sublingual region)	61	46%	10.4mo (carcinoma 4.4mo, other 10.8mo)	**Positive:** Slight, moderate or severe dysplasia, CIS, carcinoma **Negative:** No dysplasia	**Dysplasia or malignancy:** 77/101 (76%)	66	11	6	18	86%	75%	92%	62%
Jeong 2012 [10] South Korea Retrospective	Leukoplakia of lateral tongue (excluded if biopsy showed carcinoma)	54	64%	1.6 mo	**Positive:** Dysplasia, carcinoma **Negative:** No dysplasia	**Dysplasia or malignancy:** 12/15 (80%)	5	7	0	3	42%	100%	100%	30%
FN, false negative; FP, false positive; mo, months; NPV, negative predictive value; PPV, positive predictive value; Sens, sensitivity; Spec, specificity; TN, true negative; TP, true positive; yr, years.														

Sensitivity and specificity for malignant-only lesions

Four studies reported data allowing calculation of sensitivity and specificity of biopsy vs. excision for identification of malignant-only lesions (**Table 2**). The studies were conducted in Denmark [5], Italy [6], Australia [7] and South Korea [10], with prevalence rates of 7%, 61%, 67% and 85% for malignant lesions on full excision. The sensitivity of biopsy (compared to excision) for detection of malignancy was 71% and 94% in the Italian and Australian studies (not calculable in the other two studies since malignant lesions identified on biopsy were excluded). The specificity for detection of malignancy varied widely between studies (17%, 89%, 100% and 100%).

Table 2: Biopsy vs. full excision: Sensitivity and specificity for diagnosis of malignant lesions only

Author, Year **Country** **Study Design**	**Study population**	**Mean age (yr)**	**Male (%)**	**Time: biopsy to excision (mean)**	**Definition of positive/ negative case** (via reference standard)	**Prevalence: dysplasia or malignancy**	**TP**	**FN**	**FP**	**TN**	**Sens**	**Spec**	**PPV**	**NPV**
Holmstrup 2007 [5] Denmark Retrospective	Oral mucosal lesions (where biopsy showed epithelial dysplasia and/or lesions on lateral/ventral tongue or sublingual region)	61	46%	10.4mo (carcinoma 4.4mo, other 10.8mo)	**Positive:** Carcinoma **Negative:** Non-carcinoma	**Malignancy:** 7/101 (7%)	NE	7	NE	94	NE	100%	NE	93%
Jeong 2012 [10] South Korea Retrospective	Leukoplakia of lateral tongue (excluded if biopsy showed carcinoma)	54	64%	1.6 mo	**Positive:** Carcinoma **Negative:** Non-carcinoma	**Malignancy:** 10/15 (67%)	NE	10	NE	5	NE	100%	NE	33%
Pentenero 2003 [6] Italy Retrospective	Oral mucosal non-homogeneous white or white/red lesions AND final diagnosis of moderate-severe dysplasia or cancer	65	43%	<1mo	**Positive:** Carcinoma **Negative:** Non-carcinoma	**Malignancy:** 28/46 (61%)	20	8	2	16	71%	89%	91%	67%
Gnanasekaran 2015 [7] (abst) Australia Retrospective	Tongue lesions	NR	NR	NR	**Positive:** SCC **Negative:** Non-SCC	**Malignancy:** 130/153 (85%)	122	8	19	4	94%	17%	87%	33%

Abst, conference abstract; FN, false negative; FP, false positive; mo, months; NE, not estimable; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; SCC, squamous cell carcinoma; Sens, sensitivity; Spec, specificity; TN, true negative; TP, true positive; yr, years.

Concordance, under-diagnosis and over-diagnosis for dysplastic or malignant lesions

Five studies reported concordance between biopsy and excision (generally defined as the percentage of lesions where both procedures indicated the same diagnostic category for grade of dysplasia or malignancy; **Table 3**). These studies also reported rates of under-diagnosis and over-diagnosis of biopsy compared with full excision. The studies were conducted in the USA [8], Denmark [5], Italy [6], South Korea [10] and UK [9], with a prevalence of 26%, 76%, 80%, 88% and 100% respectively for dysplasia or malignancy on full excision.

Concordance between biopsy and excision across all lesions was reported as 27%, 49%, 50% and 89% across four studies [5, 8-10], while one study also reported concordance as 79% when defined as the same dysplasia category or one category different [5]. Under-diagnosis was reported in three studies as 35%, 36% and 73%, while over-diagnosis was reported in the same studies as 17%, 13% and 0% [5, 9, 10]. Concordance between biopsy and excision when considering only those lesions found to be dysplastic or malignant was 72% and 81% within two studies, while under-diagnosis rates were 24% and 17% and over-diagnosis rates were 4% and 1.4% [6, 8].

Table 3: Biopsy vs. full excision: Concordance, under-diagnosis and over-diagnosis

Author, year Country Design	Population	Mean age (yr)	Male (%)	Time: biopsy to excision (mean)	Definition of concordance	Prevalence: dysplasia or malignancy	All lesions			Dysplastic or malignant lesions		
							Concordance	Under-diagnosis	Over-diagnosis	Concordance	Under-diagnosis	Over-diagnosis
Diagnosis of dysplastic or malignant lesions												
Chen 2016 [8] USA Retrospective	Mucosal, intraosseous and salivary gland lesions (single or multiple biopsies)	47	44%	2.5mo	- Concordant if same diagnosis - Categories NR	70/272 (26%)	242/272 (89%)	NR	NR	57/70 (81%)	12/70 (17%)	1/70 (1.4%)
Holmstrup 2007 [5] Denmark Retrospective	Oral mucosal lesions (where biopsy showed epithelial dysplasia and/or lesions on lateral/ventral tongue or sublingual region)	61	46%	10.4mo (carcinoma 4.4mo, other 10.8mo)	- Concordant: a) if same diagnosis b) if same or one category different in dysplasia severity - Categories: no, slight, moderate or severe dysplasia, CIS, carcinoma	77/101 (76%)	49/101 (49%) Same or one degree different: 80/101 (79%)	35/101 (35%)	17/101 (17%)	NR	NR	NR
Jeong 2012 [10] South Korea Retrospective	Leukoplakia of lateral tongue (excluded if biopsy showed carcinoma)	54	64%	1.6 mo	- Concordant if same diagnostic category - Categories: 1) no dysplasia; 2) dysplasia; 3) carcinoma	12/15 (80%)	4/15 (27%)	11/15 (73%)	0/15 (0%)	NR	NR	NR
Pentenero 2003 [6] Italy Retrospective	Oral mucosal non-homogeneous white or white/red lesions AND final diagnosis of moderate-severe dysplasia or cancer	65	43%	<1mo	- Concordant if same diagnostic category - Categories: 1) no or mild dysplasia; 2) moderate or severe dysplasia or CIS; 3) carcinoma	46/46 (100%)	NR	NR	NR	33/46 (72%)	11/46 (24%)	2/46 (4%)
Thomson 2017 [9] UK Retrospective	Newly presenting, single oral mucosal lesions (leukoplakia, erythroleukoplakia or erythroplakia)	60 (range 23-94)	59%	1.5 - 3mo	- Concordant if same diagnostic category - Categories: no, mild, moderate or severe dysplasia, CIS, SCC	522/590 (88%)	307/609 (50%)	220/609 (36%)	82/609 (13%)	NR	NR	NR
Abst, conference abstract; CIS, carcinoma-in-situ; mo, months; NR, not reported; SCC, squamous cell carcinoma; yr, years.												

Reasons for discordance

Reasons for discordance, as noted in the included studies, are outlined in **Table 4**. The USA study (Chen et al., 2016) [8] reported reasons for discordance: of the 30 discordant cases, 18 (60%) were due to sampling error where biopsy tissue was not representative of the whole lesion; 7 (23%) were due to pathologist discordance; 4 (13%) were due to insufficient tissue in the biopsy specimen; and 1 (3%) was due to obscuring inflammation. This study also reported that concordant cases had a larger average biopsy volume than discordant cases (1.53 vs. 0.42 cm^3, p=0.063). Also, among the 12 lesions with multiple biopsies, overall concordance was 83% while both the under-diagnosis and over-diagnosis rates were 0% (the nature of the remaining 17% is not clear, but potentially these cases had differing diagnosis of similar severity).

The South Korean study (Jeong et al., 2012) [10] reported similar reasons for discordance; these included sampling error where biopsy tissue was not representative (for large lesions); superficial biopsy (frequently associated with punch biopsy); and errors in pathology specimen preparation due to small specimen size (resulting in tangential cutting affecting evaluation of the submucosal area). Numbers of discordant cases for each issue were not reported in this study.

Table 4: Biopsy vs. full excision: Reasons for discordance

Author, year Country	**Factors related to concordance**	**Reasons for discordance**
Chen 2016 [8] USA	* **Volume of biopsy:** Concordant cases larger average biopsy volume (1.53 vs. 0.42 cm^3, p=0.063) ***Multiple-site biopsies:** Lesions with multiple biopsies (n = 12): overall concordance 83%, under-diagnosis 0%, overdiagnosis 0% (note: if 1 site consistent then considered concordant)	**Reasons for discordance:** * Sampling error (biopsy tissue not representative of whole lesion) 18/30 (60%) * Pathologist discordance 7/30 (23%) * Insufficient tissue in biopsy specimen 4/30 (13%) *Obscuring inflammation 1/30 (3%)
Jeong 2012 [10] South Korea		**Reasons for discordance:** * Sampling error within large lesion * Superficial biopsy (frequently associated with punch biopsy) * Error in pathology specimen preparation due to small specimen size (resulting in tangential cutting affecting evaluation of submucosal area)

Discussion

Overall, there was surprisingly little published evidence relating to the diagnostic accuracy of biopsy of potentially malignant and malignant oral lesions compared with full excision. Only six studies undertaken within developed countries were identified, with little consistency in measures of diagnostic accuracy between studies, making comparison difficult.

The review findings can be summarized as follows: for identification of dysplasia or malignancy, sensitivity of biopsy (compared to excision) ranged from 42% to 86% and specificity from 75% to 100% in the two studies reporting this [5, 10], while measures of concordance varied from 27% to 89% across four studies [5, 8-10]. For identification of malignant-only lesions, sensitivity was 71% and 94% in the two studies where this was calculable, while specificity ranged from 17% to 100% across four studies [5-7].

While five articles reported data on concordance between biopsy and excision, definitions of concordance were not always clear and were not entirely consistent between articles, and these data could not be used to calculate sensitivity and specificity. Most of the articles defined concordance as the same grade of dysplasia, while one also assessed concordance defined as either the same grade or one grade different, and another used broader categories of no dysplasia, dysplasia or carcinoma. Whether one grade difference represents acceptable concordance may depend on the situation, for example whether the difference in grading means there is a missed opportunity for further treatment (such as excision).

The differences in accuracy across studies may have been due to factors such as prevalence of dysplastic or malignant lesions in the study sample (which ranged from 26% to 100% across studies), prevalence of malignant lesions (which ranged from 7% to 85%), nature of the lesions included (in terms of visual appearance and location in the oral cavity), country and setting of the study, and the precise nature of the biopsy and excision procedures and histological analysis (which were not well reported).

Possible reasons for discordance between biopsy and excision, as suggested within the reviewed studies, included: biopsy tissue not being representative of the whole lesion; insufficient tissue in the biopsy; superficial biopsy (often due to punch biopsy); errors in specimen preparation resulting in tangential cutting; and pathologist discordance [5, 8, 10]. It is also possible that some lesions did not

show dysplasia or carcinoma at the time of biopsy but had progressed by the time of excision [5], as time from biopsy to excision ranged from less than 1 month to a mean of 10.4 months across studies. There may also have been bias in the design of the included studies, in that patients were only included if they had undergone both biopsy and excision, thereby excluding patients undergoing biopsy only, who may not have been deemed to require full excision or whose disease may not have progressed following biopsy. In many cases, particularly in those with mild to moderate epithelial dysplasia, full excision would only rarely form part of routine clinical management, thus, the cohorts are biased towards those with higher grade dysplasia.

In terms of implications for practice, the included studies make various suggestions for improving biopsy accuracy, including: ensuring uniform biopsy depth; multiple biopsies of large lesions; sampling of the full lesion in the case of encapsulated lesions; and liaising with pathologists regarding sampling, block thickness and tissue orientation [7, 8]. Other authors have also noted the importance of taking deeper levels of sections throughout the specimen [11], as well as the issue of pathologist concordance [12]. It is also suggested that patients should receive regular follow-up of the lesion following biopsy (e.g. every 3-6 months) irrespective of whether dysplasia was identified [5].

Further research would be valuable to assess the diagnostic accuracy of biopsy compared with full excision for different patient populations and lesion types and in different clinical settings. In addition, it would be useful to evaluate approaches which may improve biopsy accuracy (e.g. biopsy and histopathology techniques, including the potential use of digital pathology and artificial intelligence[13]) or compensate for imperfect accuracy (e.g. regular follow-up).

Conclusions

A systematic review of the diagnostic accuracy of biopsy versus full excision for diagnosis of oral cancer and potentially malignant disorders identified six studies. There was substantial variation both in the measures of diagnostic outcomes reported, and in the findings regarding accuracy of biopsy compared with excision. Suggestions for improving biopsy accuracy included ensuring adequate biopsy depth; multiple biopsies of large lesions; addressing discordance between pathologists; and regular follow-up. This analysis highlights the lack of research in this area, and the wide variability in the few reported studies, which hinders firm conclusions. There is a pressing need for further research to assess biopsy accuracy for different lesion types and settings, as well as to evaluate approaches to improve biopsy accuracy.

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424.

[2] Dissanayaka WL, Pitiyage G, Kumarasiri PVR, Liyanage RLPR, Dias KD, Tilakaratne WM. Clinical and histopathologic parameters in survival of oral squamous cell carcinoma. *Oral surgery, oral medicine, oral pathology and oral radiology*. 2012;113:518-25.

[3] Sciubba JJ. Oral cancer. *American journal of clinical dermatology*. 2001;2:239-51.

[4] Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer*. 2018;18:269-82.

[5] Holmstrup P, Vedtofte P, Reibel J, Stoltze K. Oral premalignant lesions: is a biopsy reliable? *Journal of oral pathology & medicine*. 2007;36:262-6.

[6] Pentenero M, Carrozzo M, Pagano M, Galliano D, Broccoletti R, Scully C, et al. Oral mucosal dysplastic lesions and early squamous cell carcinomas: underdiagnosis from incisional biopsy. *Oral diseases*. 2003;9:68-72.

[7] Gnanasekaran T, Latis S, Iseli T, Wiesenfeld D. Are incisional tongue biopsies for squamous cell carcinoma representative? *International Journal of Oral and Maxillofacial Surgery*. 2015;44:e69.

[8] Chen S, Forman M, Sadow PM, August M. The Diagnostic Accuracy of Incisional Biopsy in the Oral Cavity. *Journal of oral and maxillofacial surgery*. 2016;74:959-64.

[9] Thomson PJ, Goodson ML, Cocks K, Turner JE. Interventional laser surgery for oral potentially malignant disorders: a longitudinal patient cohort study. *Int J Oral Maxillofac Surg*. 2017;46:337-42.

[10] Jeong WJ, Paik JH, Cho SW, Sung MW, Kim KH, Ahn SH. Excisional biopsy for management of lateral tongue leukoplakia. *J Oral Pathol Med*. 2012;41:384-8.

[11] Manyam R, Reena K, Saraswathi TR. Role of deeper sections in diagnostic oral histopathology: a retrospective study. *Indian J Dent Res*. 2011;22:62-5.

[12] Kujan O, Khattab A, Oliver RJ, Roberts SA, Thakker N, Sloan P. Why oral histopathology suffers inter-observer variability on grading oral epithelial dysplasia: an attempt to understand the sources of variation. *Oral Oncol*. 2007;43:224-31.

[13] Bashir RMS, Mahmood H, Shaban M, Raza SEA, Fraz MM. Automated grade classification of oral epithelial dysplasia using morphometric analysis of histology images. Proc SPIE 11320, Medical Imaging 2020: *Digital Pathology*. 2020.