Ploidy: Friend or Foe

EW Odell

Outline of a plenary lecture to be held at the
International Association for Oral Pathologists Biannual Meeting Brisbane 2006

There was great interest when a paper from the Norwegian Radium Hospital, published in the New England Journal of Medicine in 2001, proposed that keratinocyte DNA content predicted malignant transformation in oral lesions. Using an automated image-based ploidy analysis system the authors claimed that the positive predictive value of an aneuploid result was 84% and the negative predictive value of a diploid result was 97%. These impressive results were accompanied by data showing that traditional dysplasia grading had a very poor predictive value.

Since publication oral pathologists have been unsure whether ploidy analysis really offers significant benefits. This remains the only large scale study of the technique and the poor predictive value of the dysplasia grading is not only contrary to series published previously but also flies in the face of most diagnostic pathologists' experience. Alternative methods of assessing ploidy do not appear to have the predictive value, have not been tested against outcome or are not widely available.

Before considering ploidy analysis, it is necessary to examine the scientific basis for dysplasia grading itself, because this is the gold standard against which any new technique will be assessed. Despite its importance, the scoring of dysplasia and identification of dysplasia remain problematic areas in histopathology practice. Many publications include lists of features of dysplasia but many are imprecisely defined and there is no explanation of how their severity should be recorded and no indication of the relative importance of each. There is no supporting experimental data to support either the "thirds" grading system or the new WHO grading system. Large scale studies giving positive and negative predictive values for dysplasia grading have not been performed. Indeed, much of the recent literature is critical, denigrating oral dysplasia grading as subjective, poorly predictive and lacking in reproducibility.

To address these issues and assess the value of ploidy analysis our group has set up a large follow up study of oral potentially malignant lesions. We have identified 3255 patients with oral red and white lesions and are collecting clinical and histopathological data on them. To date we have identified 699 lesions from 416 patients that were subjected to biopsy between 1990 and 1999 and sent to the laboratory with clinical suspicion of malignant potential. Ploidy analysis has been performed on all these lesions. Dysplasia grading is available for 907 lesions from 422 patients. Follow up data for a minimum post biopsy period of 6 years was provided by Thames Cancer Registry and the Office of National Statistics. Ethical approval to perform the study without patient consent was obtained from the Secretary of State under section 60 of the Health and Social Care Act 2001.

I will use data from this study to address the issue of whether ploidy analysis should be welcomed or rejected. Some raise objections to ploidy analysis on grounds of cost or fear that diagnosis may be transferred to non-pathologists without the ability to interpret clinical information. However, our response to ploidy should be determined by only one issue, its effectiveness.

The concept that DNA aneuploidy precedes carcinoma has a sound scientific basis. Chromosomal instability is thought to be a fundamental process in neoplasia. Loss of control of chromosomal complement is a progressive change that can result from mutation, inactivation or over expression of many oncogenes such as those controlling the cell cycle checkpoints, spindles and segregation genes and DNA mismatch repair genes. Aneuploidy has a genome-wide effect on the transcriptome. Oral lesions screened by comparative genomic hybridisation and SNP chip CGH array analysis have been shown to harbour chromosomal imbalances at numerous sites including, 1p, 3p, 3q, 8q, 9p, 11q ,15q, 17p, 21q. It appears that all dysplastic lesions harbour small deletions and duplications indicating chromosomal instability at the molecular level.

Studies indicating a possible predictive value, in addition to the Norwegian study, include image-based analysis of brush biopsy samples from Germany and of biopsy material from Sweden. These will be reviewed.

There are two issues that need to be addressed before coming to the conclusion that detection of DNA aneuploidy predicts transformation. The first is whether these anomalies are truly oncogenic or reflect non-specific effects of chromosomal instability. It is possible that cells with a range of microdeletions and duplications may be those least able to become effective malignant cells, because independent growth, invasion and metastasis will require the coordinated expression and control of many genes. The second is, if predictive, at what detection level the aneuploidy should be detected to be predictive. Just because DNA aneuploidy precedes malignant transformation does not mean that it will be an accurate predictor.

DNA ploidy analysis in general suffers problems of interpretation and technique. In particular, there are occasional aneuploid cells in normal individuals and these small populations must be ignored. This is achieved using threshold values in the diagnostic criteria so that the technique risks an integral false negative rate. Any diagnostic threshold needs to be determined experimentally to define the resulting predictive value of the test. This does not appear to have been performed for oral lesions.

In our studies we have used the Fairfield ploidy analysis system, with the automated version of the software used in the original Norwegian trial. We have found that analysis may be readily performed on fixed paraffin-embedded tissue from routine biopsy samples. Oral mucosa is particularly difficult to extract nuclei from and the coefficient of variation of the diploid peak, while usually less than 5%, sometimes exceeds this value. This suggests that the threshold of detection is approximately 1% total DNA content per cell. All normal and hyperplastic lesions we have tested have been diploid.

The lecture will include a brief review of the technique, with explanations of the control experiments required by new users to validate the technique in their own populations. This will also include a review of the clinical and laboratory problems that must be overcome in setting up a new service. The problem of sampling error will receive particular attention with data from our publications on oral carcinoma.

At the time of writing, the data analysis for our study is incomplete because the follow up for some patients is awaited from the Office of National Statistics. In this synopsis I will provide illustrative data only, based on whole patient rather than lesion analysis for 1096 lesions. The results will be finalised at the end of April 2006 when the follow-up period ends and site analysis will be possible. The figures in the following paragraphs should be regarded as indicative only and are not for further dissemination.

As noted above, there has been no study that published predictive values for dysplasia grading by conventional means. In our study the following cases were identified for which a ploidy result could be obtained:

Dysplasia grade	**None**	**Mild**	**Moderate**	**Severe**
Total	509	212	215	149
no transformed	0	32	32	23

If this data is analysed on the basis of the highest grade experienced by any one patient the sensitivity approximates 40% and the specificity 8-16%, corresponding to the following predictive values for malignant transformation:

	Positive predictive value %	**Negative predictive value %**
No dysplasia		100
At worst mild	4.9	
At worst moderate	10.4	3.4
At worst severe	16.4	2.8

A ploidy result is available for 380 lesions at time of writing. When analyzed in an identical patient-based analysis the corresponding predictive values are

	Positive predictive value %	**Negative predictive value %**
Diploid	2.7	
At worst tetraploid	1.6	2.7
At worst aneuploid	14	3.1

This analysis suggests that a DNA ploidy result of aneuploidy has approximately the same predictive value as severe dysplasia, however, when site is taken into account these values may separate. Some published data suggests that aneuploidy reflects field change and so this site specific data will be of great interest. Because the sensitivity and specificity of the two techniques vary, they appear to be identifying different subpopulations.

One problem interpreting this data is the difficulty of categorising ploidy histograms. The range of ploidy anomalies is wide and the threshold values for tetraploid and aneuploid diagnoses have been determined on prostate gland and other tissues. A further 59 lesions were categorised as suspicious and only 1 patient developed an oral carcinoma, suggesting that these cases do not differ significantly from diploid lesions.

Only one mild dysplasia was diagnosed as aneuploid.

The results of our study appear to contradict the previously published data from the Norwegian study. However, there are significant differences between the study population and analysis. In our study a wider range of lesions has been sampled and the clinical presentations suggest a much lower risk population. In the Norwegian study there was a very high proportion of lesions with moderate (40%) and severe (30%) dysplasia and the study population suffered a 24% overall transformation rate during follow up from 1-13 years, much higher than in most routine referral treatment centres. It is to be expected that the positive predictive value of any test will fall if it is applied to a population in which the outcome is less prevalent. We would not have expected to find a predictive value as high as that previously reported.

A challenge therefore remains: to determine whether the predictive value can be increased by better patient or lesion selection for ploidy analysis. We are also undertaking receiver-operator-characteristic analysis to identify more specific cut off values for the continuous parameters used to categorise ploidy histograms.

Data will also be presented on specific subgroups of dysplastic lesions in which dysplasia grading is recognised to be difficult or misleading: lichen planus and proliferative verrucous leukoplakia

It remains unclear whether lichen planus has a risk of malignant transformation or whether some dysplastic lesions with a prominent host response are indistinguishable from lichen planus clinically and histopathologically. It is clear that identifying mild degrees of dysplasia in a background of florid host response is extremely difficult and it is possible that ploidy analysis might be useful. We have compared 13 cases that meet strict criteria for possible malignant transformation in lichen planus and compared their ploidy status with 80 samples from typical bilateral reticular lesions of lichen planus from patients who did not take potentially causative drugs and were non smokers. A second control group included samples from 50 patients presenting with clinically atypical lichen planus, defined as cases in which the clinical appearance might suggest a transformation risk: unilateral lesions, lesions affecting the floor of mouth or other high risk sites and unusual lesions in smokers. All such lesions analysed to time of writing have proved diploid.

Of the 13 cases with transformation, 8 had previous samples and aneuploidy preceded carcinoma in 3 and tetraploidy in 1. In 2 of the remaining 4 no ploidy anomaly was detected but the small carcinomas were also diploid. This suggests that the positive predictive value in lichen planus is high.

Verrucous lesions also present particular problems because their relatively low incidence of cytological atypia is not commensurate with the fact that their risk of transformation is higher than flat lesions. To address whether ploidy analysis might help in this situation we have evaluated the technique in proliferative verrucous leukoplakia (PVL). The clinical and pathological features of 6 patients with PVL were reviewed and all their biopsy specimens were subjected to image-based DNA ploidy analysis. Most patients were female and the average age on first biopsy was 66 years. Only one patient reported both tobacco smoking and alcohol intake. The most frequently affected sites were alveolar ridge and/or gingiva (6/6), buccal mucosa (3/6), palate (3/6), tongue (2/6), buccal sulcus (2/6), and lip (1/6). All patients progressed to carcinoma and three developed multiple primary carcinomas, either invasive or verrucous.

Five patients showed aneuploidy in lesions prior to diagnosis of carcinoma. If the same results had been obtained prospectively, a ploidy anomaly would have predicted malignant transformation in 4 cases and probably in a fifth for whom DNA ploidy failed to meet diagnostic criteria but was suspicious of aneuploidy. This is a higher predictive value than in our larger study and particularly interesting in view of the low level of nuclear atypia in these lesions. However, intriguingly, dysplasia also predicted transformation equally accurately, perhaps reflecting our tendency to upgrade histological features in verrucous lesions.

The lecture will conclude with a short analysis of cost benefit and practical problems introducing a diagnostic service. We have found that many biopsy specimens are too small for diagnosis and an element of operator education is required to ensure that specimens can be analysed. Fixation in unbuffered formalin produces a higher CV than buffered formalin. This may render retrospective studies necessary to quality assure the technique difficult. No formal cost benefit analysis has been performed but one Biomedical Scientist can process approximately 30-40 specimens each week. Up to 15% of these may require re-preparation, particularly while developing the necessary technical skills. It takes one pathologist between one half and one and a half hours to check the cytology and make the diagnoses on 10 cases, depending on the quality of the preparation and number of nuclei sampled.

Even if predictive, it remains to be established that ploidy can be integrated into care pathways with a positive cost benefit. Only if clinicians can exclude diploid lesions from follow-up will the technique be economic overall, even though it may be significant in reducing morbidity and mortality.

Are there alternatives to this time consuming and costly technique? The search for predictive molecular and immunocytochemical markers has so far proved largely fruitless. A number of markers show correlation with dysplasia, particularly with basaloid compartment expansion. However, none has yet proved predictive. We have tested a range of markers, COX2, p53 (Do7), cytokeratin 8, cytokeratin 18, p16INK4a and Ki67 (Mib1) against 20 oral lesions, 7 of which were from patients what subsequently developed oral carcinoma. Ten lesions were diploid and 10 aneuploid. No marker showed a correlation with DNA ploidy status suggesting that aneuploidy is too complex a state to be identified by small panels of markers. This data is consistent with one previous report for p53 and p21 from Sweden.

In conclusion, in our series ploidy has proved a valid technique for predicting malignant transformation of oral lesions. However, it does not appear to hold great advantages over traditional methods of dysplasia grading when used in a primary diagnostic role for all clinically suspicious lesions. It is able to predict transformation in verrucous lesions without obvious nuclear atypia and in lesions with a prominent host response such as those resembling lichen planus. In a minority of cases aneuploidy is revealed unexpectedly and we continue to work to refine selection criteria for ploidy analysis as application to improve the cost:benefit ratio. These further analyses will be presented in the talk.

The talk will conclude with thanks to members of the research group and funding agencies.

Postscript: Since this presentation, this work has shown a greater benefit with ploidy detecting transformation risk in non-dysplastic and mildly dysplastic lesions and its use has become routine in the authors centre. Subsequently the Norweigan data was withdrawn. The final research outcome from this research took many years to complete and is now available at: *doi:10.1158/1940-6207.CAPR-13-0001*