Dear Sir,

Oral cancer and precancer screening studies have high discriminatory ability. Evidence-Based Dentistry 2000; 3:79–80

In your 'further discussion' relating to our paper¹ and the commentary by Farsai and Henshaw, we were pleased to note your opening statement, "As we endeavour where possible to allow the original author to see the commentary on their article prior to publication, this leads to interesting interchanges." You did kindly allow us to see and respond both to the commentary and the draft of your discussion, and in response we sought to address a number of important criticisms of our paper made by the commentators. We were thus a little disappointed that our response was not included in your commentary, and felt that several key points set down in defence of our work were, unfortunately, either misconstrued or overlooked in the further discussion as published. As a result the interchange was perhaps not as interesting, or indeed relevant, as it might have been. In order to redress what we believe still stand as unanswered, unjustified, criticisms of our article, we would be grateful if you would allow us to reproduce here, in our own words, those particular parts of our response that we believe did not receive a full and adequate airing. We will refer only to what we regard as the most significant areas of contention.

First, in response to the initial draft discussion, we said that the aspect with which we most wished to take further issue related to the question of the 'gold standard' against which oral cancer screening tests were assessed. The screen is intended to detect relevant lesions in apparently healthy people that may be cancer or precancer. The appropriate gold standard, albeit a 'soft' one, is independent diagnosis by an expert. We believe that it cannot be stressed too highly that biopsy (the gold standard of choice according to the commentators) is not an option in the validation of population screening tests for oral cancer and precancer, where sensitivity and specificity are the outcome measures, as is standard practice. We are not sure that the commentators appreciate how these measures are computed. Perhaps they are not acquainted with the classic monograph on screening by Wilson and Jungner;² they seem to imply that it is only necessary to apply the validating criterion to the subjects who test positive, whereas, of course, it has to be applied to all subjects — positive and negative — or at least to a random selection of the negatives in addition to the positives. In your commentary, you only address the possible ethical problems of subjecting apparently healthy mucosa to biopsy. We do not think the ethics of biopsy of clinically normal mucosa is a relevant issue; if it was necessary or of use it could be done. It is in fact impossible

for a biopsy of normal mucosa to be of any use or relevance in this context. Put simply, where would one take the biopsy? Would it be taken from the mucosa of the cheek, base of the tongue or oropharynx, for example? In order to avoid perpetuation of the misconception about the assessment of the performance of screening tests for oral cancer and precancer, we believe it would have been helpful if you had added a much stronger emphasis to this point in your discussion.

Our second point concerns the extent to which the conclusions we drew from our work were supported by the results. In our original report we stated that, "the meta-analysis indicated that overall the studies had discriminatory ability" and that, "the discriminatory ability demonstrated by the screening personnel, irrespective of their grade and training, was generally of a fairly high order and did not appear to be greatly affected by the varied circumstances of the studies". Both these statements, in relation to the studies ultimately included, are fully supported by the findings of the meta-analysis. Yet according to Farsai and Henshaw, the article contained inadequate information to fully evaluate this conclusion. We disagree most strongly with this view. In your commentary you dwelled on this criticism but overlooked Farsai and Henshaw's other point, that, "It is important to note, however, that the considerable heterogeneity only moderately influenced the sensitivity and specificity values." This view supports our statements above from the original report and is one reason, as well as lack of space, that we did not describe in detail the varied circumstances of each study. Indeed, under the circumstances, we concur with your view that, "The more traditional approach is for the interested reader to read the original papers." At the same time we would stress that we did not, and indeed would not, generalise from our results to suggest that intensive training and calibration of screeners is not necessary. We have in fact stressed the importance of this in many of our group's other publications.

David R Moles, Martin C Downer, Paul M Speight
London, UK

1. Moles DR, Downer MC, Speight PM. Meta-analysis of measures of performance reported in oral cancer and precancer screening studies. *Br Dent J* 2002; 192:340–344.
2. Wilson JMG, Junger G. Principles and Practice of Screening for Disease. Geneva: World Health Organisation: 1968.

Evidence-Based Dentistry (2003) **4**, 45.
doi:10.1038/sj.ebd.6400187