More Accurate Oral Cancer Screening With Fewer Salivary Biomarkers

Biomarkers in Cancer
Volume 9: 1–15

DOI: 10.1177/1179299X17732007

James Michael Menke[1], Md Shahidul Ahsan[2] and Suan Phaik Khoo[3]

[1]A.T. Still Research Institute, A.T. Still University, Mesa, AZ, USA. [2]Department of Oral Pathology, Radiology and Medicine, College of Dentistry and Dental Clinics, The University of Iowa, Iowa City, IA, USA. [3]Department of Oral Diagnostic and Surgical Sciences, School of Dentistry, International Medical University (IMU), Kuala Lumpur, Malaysia.

ABSTRACT: Signal detection and Bayesian inferential tools were applied to salivary biomarkers to improve screening accuracy and efficiency in detecting oral squamous cell carcinoma (OSCC). Potential cancer biomarkers are identified by significant differences in assay concentrations, receiver operating characteristic areas under the curve (AUCs), sensitivity, and specificity. However, the end goal is to report to individual patients their risk of having disease given positive or negative test results. Likelihood ratios (LRs) and Bayes factors (BFs) estimate evidential support and compile biomarker information to optimize screening accuracy. In total, 26 of 77 biomarkers were mentioned as having been tested at least twice in 137 studies and published in 16 summary papers through 2014. Studies represented 10 212 OSCC and 25 645 healthy patients. The measure of biomarker and panel information value was number of biomarkers needed to approximate 100% positive predictive value (PPV). As few as 5 biomarkers could achieve nearly 100% PPV for a disease prevalence of 0.2% when biomarkers were ordered from highest to lowest LR. When sequentially interpreting biomarker tests, high specificity was more important than test sensitivity in achieving rapid convergence toward a high PPV. Biomarkers ranked from highest to lowest LR were more informative and easier to interpret than AUC or Youden index. The proposed method should be applied to more recently published biomarker data to test its screening value.

KEYWORDS: evidence synthesis, cancer, biomarkers, likelihood ratio, signal detection, Bayesian analysis

RECEIVED: June 13, 2017. **ACCEPTED:** August 21, 2017.

PEER REVIEW: Two peer reviewers contributed to the peer review report. Reviewers' reports totaled 417 words, excluding any confidential comments to the academic editor.

TYPE: Technical Advance

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was initiated with generous support from the International Medical University research grant no. IMUJC 290813 (71st meeting), project ID no. IMU 286/2013. The authors are very grateful for International Medical University's support of this project.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: James Michael Menke, A.T. Still Research Institute, A.T. Still University, 5850 East Still Circle, Mesa, AZ 85206, USA. Email: jmenke@atsu.edu

Introduction

Purpose

The purpose of this study was to develop methods for improving salivary biomarker efficiencies in oral cancer screening with inferential tools adapted from clinical epidemiology and Bayesian analysis.

Problem

Even as screening and diagnostic technologies drop in price, methodological efficiencies can improve access and hasten implementation of salivary disease screening. Reducing test complexity, cost, and interpretation of results while increasing test information and robustness is a useful goal of any medical screening.

However, the current methods of biomarker discovery might limit their value in detecting disease. Potential salivary biomarkers are first observed as different concentrations in the saliva of healthy patients versus that of patients with cancer. From these samples, some or all of 4 attributes are commonly reported: sensitivities, specificities, areas under the curve (AUCs), and statistical significance in biomarker concentration between healthy and cancer groups.

However, from a clinical or policy decision's perspective, more information is needed than just these 4 statistics. First, sensitivity and specificity are test characteristics and not related to patient health status. Statistical significance is subject to misinterpretation and often fails to meet underlying assumptions.[1] Finally, AUC is not easily computed and interpreted for use in medical decisions. Area under the curve may represent biomarker accuracy in a population but does not offer an estimate of personal risk given positive test results. However, the positive predictive value (PPV) includes test accuracy characteristics, accounts for disease prevalence, and states personal patient risk as a probability of having a disease.

Oral cancer

Oral cancer is a significant disease worldwide with up to 400 000 new cases and almost 130 000 deaths annually.[2] About 90% of oral cancers are oral squamous cell carcinoma (OSCC), and 80% occur in Southeast Asian countries. In the United States, the 5-year OSCC survival rate is 60%, but higher incidence and lower survival rates have been reported in some South Asian countries.[3–5] Most OSCC lesions are small, asymptomatic, and easily overlooked or misjudged making early detection a challenge. Risk factors include alcohol and tobacco consumption, use of betel quid chew, and alcohol-based mouthwashes.[6] More than 95% of those with OSCC smoke tobacco, drink alcohol, or both. In recent decades, more cases of OSCC are emerging without known attributable risk factors.[7]

Early detection of OSCC would improve survival, mortality, and morbidity. The gold standard for diagnosis is a biopsy of the suspicious lesion, but this test is not convenient for screening purposes because of its invasive nature, high cost, and need for specially trained medical personnel and equipment. However, saliva sampling to assess disease biomarkers is simple and less expensive, and multiple samples can be obtained for disease monitoring. Research suggests that salivary biomarkers have high discriminatory power for multiple diseases and cancers,[8–12] and they have been well characterized and validated by comparisons of biomarker levels in healthy and cancerous oral epithelia. Development from normal to OSCC cells leads to altered expression of proteins and messenger RNA markers in saliva. As such, transcriptomic and proteomic approaches to biomarker development have had the most success in clinical testing.[13]

The most informative screening tests are both highly sensitive and specific.[14] High *S*ensitivity with a *N*egative finding rules out disease, the SNout mnemonic. A highly *S*pecific test with a *P*ositive will rule-in disease, the SPin mnemonic. It may seem intuitive that more tests improve screening accuracy, but they do not, unless read as conditional probabilities where each biomarker test is read in context with others.

Independence of biomarker probabilities can only be incorrect because they are derived from the same biological system (person). For instance, 2 truly independent and positive cancer biomarkers (designated T+) in just 3% and 4% in healthy people would co-occur at just 0.12% (3% × 4%) in healthy individuals. The simultaneous occurrence of such rare biomarkers is prima facie support of a nonhealthy state, but such evidence neglects disease prevalence and does not indicate the personal risk of disease for the screened individual with both biomarkers.

Likelihood ratios reveal biomarker information values

The likelihood ratio (LR) describes a medical test's information value. A biomarker's ratio of true hit rate to false positive rate, is a positive LR. LR is used in signal detection research,[14,15] with analogous application to cancer detection. The LR is an index of how well a biomarker can separate the signal from noise.

Each biomarker offers information about the presence or absence of a disease state with varying degrees of accuracy. This task is similar to that of psychological testing, where each test item reflects the extent of an underlying construct, generally called θ. With cancer as θ, biomarker LRs can improve screening accuracy with fewer biomarker items, i.e., improve screening efficiency.

Bayes factors as evidential support

Bayes factors (BF) estimate a biomarker's strength of evidence or evidential support (ES).[15–17] Mathematically, the BF is the ratio of probability of data occurring with and without a specific biomarker. Evidential support, the natural logarithm of the BF, expresses the contribution to screening information provided by the biomarker.[15]

Correcting for inflated personal risk, PPV

Figure 1 illustrates the effects of prevalence on PPV under 6 hypothetical biomarkers, each of a different fixed sensitivity and specificity. The biomarker on the red line has 90% sensitivity and 90% specificity. Other lines in Figure 1 represent different combinations of sensitivity and specificity. In all cases, PPV increases with prevalence. For prevalence under 20% (an exaggerated maximum for disease prevalence), the 6 combinations of sensitivities and specificities show wide discrepancies in PPV. Therefore, all retrospective case-control studies in which 50% of cases have cancer and the other 50% are without cancer (and presumably healthy) inflate PPV estimates. In terms of sensitivity and specificity, some degree of each are required to improve up on prevalence as the a priori best estimate of personal cancer status.

Materials and Methods

Studies chosen

The authors selected 137 analyses with keywords—biomarkers, oral cancer, oral squamous cell carcinoma, sensitivity, and specificity—in 16 papers from a total of 80 revie wed.[3,4,7,11–13,16–86] All studies were chosen from searches in PubMed and Web of Knowledge databases in 2014. Accepted papers reported at least 2 separate estimates of sensitivity, specificity, AUC, or statistical significances between diseased and healthy groups for each biomarker. Two biomarker parameter estimates allow some degree of reliability for analysis. Twenty-six biomarkers met the criteria.

The final 26 selected biomarkers and associated panels are summarized in Appendix 1. Outcomes were transferred to Meta-DiSc and R language for computation of LRs. Biomarkers selected were: absent in melanoma 1 (AIM1), calcitonin-related polypeptide alpha (CALCA), cyclin-A1 (CCNA1), cyclin-D2 (CCND2), cadherin-1 (CDH1), death-associated protein (DAPK), deleted in colorectal carcinoma (DCC), erythrocyte sedimentation rate (ESR), hypermethylated in cancer 1 (HIC1), homeobox protein Hox-A9 (HOXA9), interleukin 1B (IL1B), interleukin 8 (IL8), O-6-methylguanine-DNA methyltransferase (MGMT), munc-18-interacting 1 (MINT1), munc-18-interacting 31 (MINT3), miR31, OAX1, p16, protein gene product 9.5 (PGP9.5), retinoic acid receptor beta (RARB), Ras association domain-containing protein 1 (RASSF1A), RNA-binding motif protein 6 (RBM6), retinoblastoma-interacting zinc-finger protein 1 (RIZ1), S100A2, sulfate adenylyltransferase (SAT), transforming growth factor beta receptor 2 (TBFBR2), and metalloproteinase inhibitor 3 (TIMP3).

Evidential support and likelihood ratio calculations

Probabilities of data under the biomarker absent (or "null") condition over the biomarker present probability of data were obtained via an hierarchical data analytic approach described elsewhere.[89,90] The ratios are Bayes Factors; the logarithms of

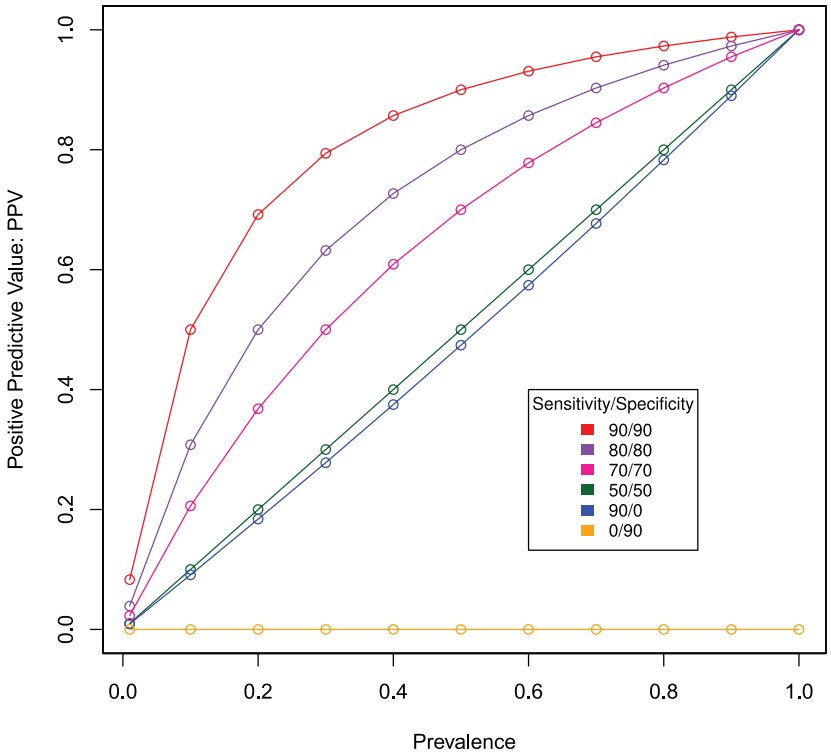

Figure 1. The effects of prevalence on positive predictive value (PPV): simulations comparing 6 sensitivity and specificity combinations.

these ratios express the ES's. In a sense, the BF is a "between" condition likelihood ratio, addressing how much more information is provided by a biomarker's presence versus its absence.

The complement is a "within" index, the LR describes biomarker accuracy or efficiency. LR is the ratio of true positives to false positives. LR times prevalence in odds form yields a PPV in odds form. LR was then adjusted for sample prevalence in case-control studies by dividing by prevalence to obtain an unbiased LR. Biomarker diagnostic accuracy was visualized using the "mada" package in R as shown in Figure 2.[91-93]

Results

Table 1 presents the 26 selected biomarkers from the 16 accepted studies listed individually and in panels with associated sensitivities (Se), specificities (Sp), AUCs, or statistical significance in biomarker amount between healthy individuals and patients with OSCC. The 5 most sensitive biomarkers were CALCA (Se = 1.00), RBM6 (0.97), S100A2 (0.96), HOXA9 (0.81), and IL8 (0.74). The 5 most specific biomarkers were RASSF1A (Sp = 0.993), AIM1 (0.991), ESR (0.986), p16 (0.950), and DCC (also 0.950). In Figure 2 (scatterplot), all OSCC biomarkers and panels reviewed in the current evidence synthesis are plotted as AUCs for detecting OSCC. The scatterplot in Figure 2A, based on 137 data points of 26 biomarkers, has an AUC of 0.760. The 9 most informative LR biomarkers in Figure 2B have an AUC of 0.818.

Information value (LR) and PPV

The 5 highest LRs from reviewed studies were ESR (LR = 7.66), AIM1 (6.98), DCC (6.20), RASSFIA (6.07), and RARB (5.38). When reading biomarkers conditionally from lowest to highest LR that have prevalence levels of 2%, 0.2%, 0.02%, 0.002%, 0.0002%, the number of biomarkers needed to achieve a PPV 100% were 11, 14, 15, 16, 18, and 20, respectively (see Figure 3). In Figure 3, the dotted line shows the number of biomarkers plotted against the PPV on the *y* axis using conditional probabilities, with green area representing 95% confidence bands needed to screen for OSCC. The blue area and long dashed line represent the Youden index plot and the pink area and solid line represent the AUC analysis. Both the Youden and AUC plots are the number of biomarkers required to approach a PPV of 100%.

By ordering from highest to lowest LR, the required number of biomarkers required to approach a PPV 100% for prevalence levels of 2%, 0.2%, 0.02%, 0.002%, and 0.0002% were 5, 8, 9, 11 and 12, respectively. Area under the curve appears to be more efficient than the LR conditional approach because it nears a PPV of 100% by starting with a high LR biomarker. For all sample prevalence conditions, 8 fewer biomarkers were needed to achieve the same OSCC screening accuracy. When the 4 least informative LR biomarkers are omitted from analysis, the number of biomarkers needed to achieve a PPV near 1.00 drops further to 3, 5, 7, 8, and 10 for each prevalence level.

The AUC and Youden index do not improve PPV by adding biomarkers. The Youden index is asymptotic at PPV = 25%, representing uncertainty for 75% of patients screened. Area under the curve was similar by increasing uncertainty: maximum screening accuracy achieved as PPV was 75%, leaving 25% error or uncertainty.

Table 1. Summary accuracy statistics for 26 oral salivary biomarkers.

BIOMARKERS	NAME	REFERENCES	PATIENTS WITH OSCC	PATIENTS WITHOUT OSCC	TOTAL NO. OF PATIENTS	PREVALENCE IN CASE-CONTROL	NO. OF COMPARISONS	SENSITIVITY	SPECIFICITY	AUC	LR (95% CI)	EVIDENTIAL SUPPORT	EVIDENCE INTER PRETATION	LR+ RANK
AIM1	Absent in melanoma	Carvalho et al,[63]	187	228	415	0.451	4	0.203 (0.174–0.233)	0.991 (0.985–0.997)	0.945	6.98 (2.81–17.31)	−3.361	Strong	2
CALCA	Calcitonin-related polypeptide α	Carvalho et al,[63]	70	50	120	0.592	2	1.000 (1.000–1.000)	0.140 (0.091–0.189)	0.976	1.17 (0.91–1.50)	−3.191	Strong	22
CCNA1	Cyclin-AI	Carvalho et al,[33] Carvalho et al,[63] Demokan et al,[50]	3811	9264	13 075	0.291	25	0.304 (0.297–0.312)	0.942 (0.939–0.944)	0.737	5.13 (4.31–6.12)	−1.956	Moderate	6
CCND2	Cyclin-D2	Carvalho et al,[63]	621	2357	2978	0.209	13	0.475 (0.455–0.495)	0.756 (0.747–0.765)	0.734	2.36 (1.77–3.16)	−1.031	Weak	16
CDH1	Cadherin-1	Carvalho et al,[63]	451	1773	2224	0.203	8	0.692 (0.670–0.714)	0.561 (0.549–0.573)	0.624	1.43 (1.07–1.91)	−1.055	Weak	21
DAPK	Death-associated protein	Carvalho et al,[33] Carvalho et al,[63]	2281	5605	7886	0.289	13	0.320 (0.310–0.329)	0.933 (0.930–0.936)	0.903	4.61 (3.73–5.69)	−1.880	Moderate	9
DCC	Deleted in colorectal carcinoma	Carvalho et al,[33] Carvalho et al,[63]	2438	6009	8447	0.289	15	0.334 (0.324–0.343)	0.950 (0.947–0.952)	0.880	6.20 (4.61–8.33)	−2.245	Moderate	3
ESR	Erythrocyte sedimentation rate	Carvalho et al,[63]	119	278	397	0.300	4	0.370 (0.325–0.414)	0.986 (0.978–0.993)	0.975	7.66 (2.15–27.34)	−3.694	Strong	1
HIC1	Hypermethylated in cancer 1	Carvalho et al,[63]	1007	4455	5462	0.184	22	0.602 (0.586–0.617)	0.730 (0.724–0.737)	0.690	2.19 (1.77–2.70)	−1.410	Weak	18
HOXA9	Homeobox protein Hox-A9	Guererro-Preston et al,[67]	16	16	32	0.500	4	0.813 (0.715–0.910)	0.875 (0.792–0.958)	0.750	2.28 (1.10–4.69)	−3.412	Strong	17
IL1B	Interleukin 1B	Brinkmann et al,[12] Brinkmann et al,[62] Elashoff et al,[74] Li et al,[22]	361	469	830	0.435	11	0.723 (0.699–0.747)	0.797 (0.779–0.816)	0.796	3.09 (2.43–3.92)	−2.330	Moderate	13
IL8	Interleukin 8	Brinkmann et al,[12] Brinkmann et al,[62] Elashoff et al,[74] Li et al,[22] St John et al,[23]	459	619	1078	0.426	14	0.736 (0.716–0.757)	0.787 (0.770–0.803)	0.832	3.27 (2.49–4.29)	−2.333	Moderate	12
MGMT	0-6-methylguanine-DNA methyltransferase	Carvalho et al,[63]	795	1438	2233	0.356	6	0.297 (0.281–0.313)	0.917 (0.909–0.924)	0.792	3.32 (2.73–4.05)	−1.534	Moderate	11
MINT1	Munc-18-interacting 1	Carvalho et al,[63]	780	1679	2459	0.317	7	0.400 (0.382–0.418)	0.887 (0.869–0.885)	0.832	4.11 (1.94–8.71)	−1.562	Moderate	10

Table 1. (Continued)

BIOMARKERS	NAME	REFERENCES	PATIENTS WITH OSCC	PATIENTS WITHOUT OSCC	TOTAL NO. OF PATIENTS	PREVALENCE IN CASE-CONTROL	NO. OF COMPARISONS	SENSITIVITY	SPECIFICITY	AUC	LR (95% CI)	EVIDENTIAL SUPPORT	EVIDENCE INTER PRETATION	LR+ RANK
MINT31	Munc-18-interacting 31	Carvalho et al,[33] Carvalho et al,[63] Demokan et al,[50]	2596	6000	8596	0.302	17	0.285 (0.276–0.294)	0.949 (0.947–0.952)	0.658	5.10 (4.15–6.28)	–2.009	Moderate	7
miR31	miR31	Liu et al,[78]	55	48	103	0.534	2	0.655 (0.590–0.719)	0.333 (0.265–0.401)	0.172	0.49 (0.01–21.21)	0.054	No evidence	27
OAZ1	Ornithine decarboxylase antizyme 1	Elashoff et al,[74] Li et al,[22]												
p16	p16	Carvalho et al,[33] Carvalho et al,[63] Demokan at al,[50]	2790	6560	9350	0.298	18	0.258 (0.250–0.266)	0.950 (0.947–0.953)	0.627	4.68 (4.07–5.37)	–1.892	Moderate	8
PGP9.5	Protein gene product 9.5	Carvalho et al,[63]	630	2516	3146	0.200	15	0.640 (0.622–0.660)	0.654 (0.645–0.664)	0.725	1.91 (1.55–2.35)	–1.218	Weak	19
RARB	Retinoic acid receptor β	Carvalho et al,[63]	101	205	306	0.330	3	0.693 (0.647–0.739)	0.917 (0.898–0.936)	0.784	5.38 (0.33–88.33)	–3.218	Strong	5
RASSF1A	Ras association domain-containing protein 1	Carvalho et al,[63]	88	139	227	0.388	2	0.114 (0.80–0.147)	0.993 (0.986–1.000)	0.141	6.07 (1.74–21.13)	–2.873	Strong	4
RBM6	RNA-binding motif protein 6	Carvalho et al,[63]	88	53	141	0.624	2	0.977 (0.961–0.993)	0.075 (0.039–0.112)	0.957	1.10 (0.77–1.58)	–1.256	Weak	23
RIZ1	Retinoblastoma-interacting zinc-finger protein 1	Carvalho et al,[63]	88	53	141	0.624	2	0.064 (0.041–0.095)	0.925 (0.888–0.961)	0.161	0.83 (0.28–2.42)	0.109	No evidence	26
S100A2	S100A2	Carvalho et al,[63]	88	47	135	0.652	2	0.955 (0.932–0.977)	0.128 (0.079–0.176)	0.617	1.10 (0.97–1.26)	–1.123	Weak	24
SAT	Sulfate adenylyltransferase	Carvalho et al,[63]	369	482	851	0.434	11	0.726 (0.703–0.749)	0.768 (0.748–0.787)	0.617	2.97 (2.23–3.96)	–2.171	Moderate	14
TGFBR2	Transforming growth factor β receptor 2	Carvalho et al,[63]	367	1081	1448	0.253	10	0.589 (0.563–0.614)	0.800 (0.788–0.812)	0.682	1.87 (1.39–2.52)	–1.745	Moderate	20
TIMP3	Metalloproteinase inhibitor 3	Carvalho et al,[63]	972	3542	4514	0.215	15	0.555 (0.539–0.570)	0.761 (0.753–0.768)	0.735	2.53 (1.69–3.26)	–1.375	Weak	15

OAZ1 was used only in combination with other biomarkers, so unique value of its contribution could not be calculated.

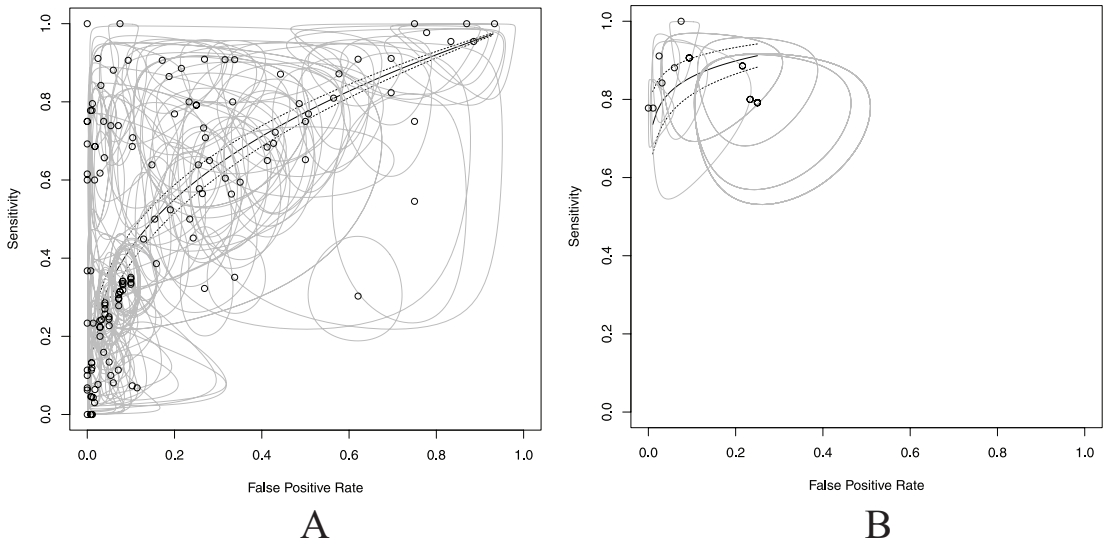

Figure 2. Biomarkers equal better test efficiencies with fewer versus more biomarkers—illustration that fewer biomarkers can have a higher positive predictive value and decreased error variance. Fewer number of biomarkers on the right (B) have a higher discriminating ability measured in area under the curve, AUC = 0.818, than the more complex "scattershot" approach on the left (A), where the AUC = 0.760.

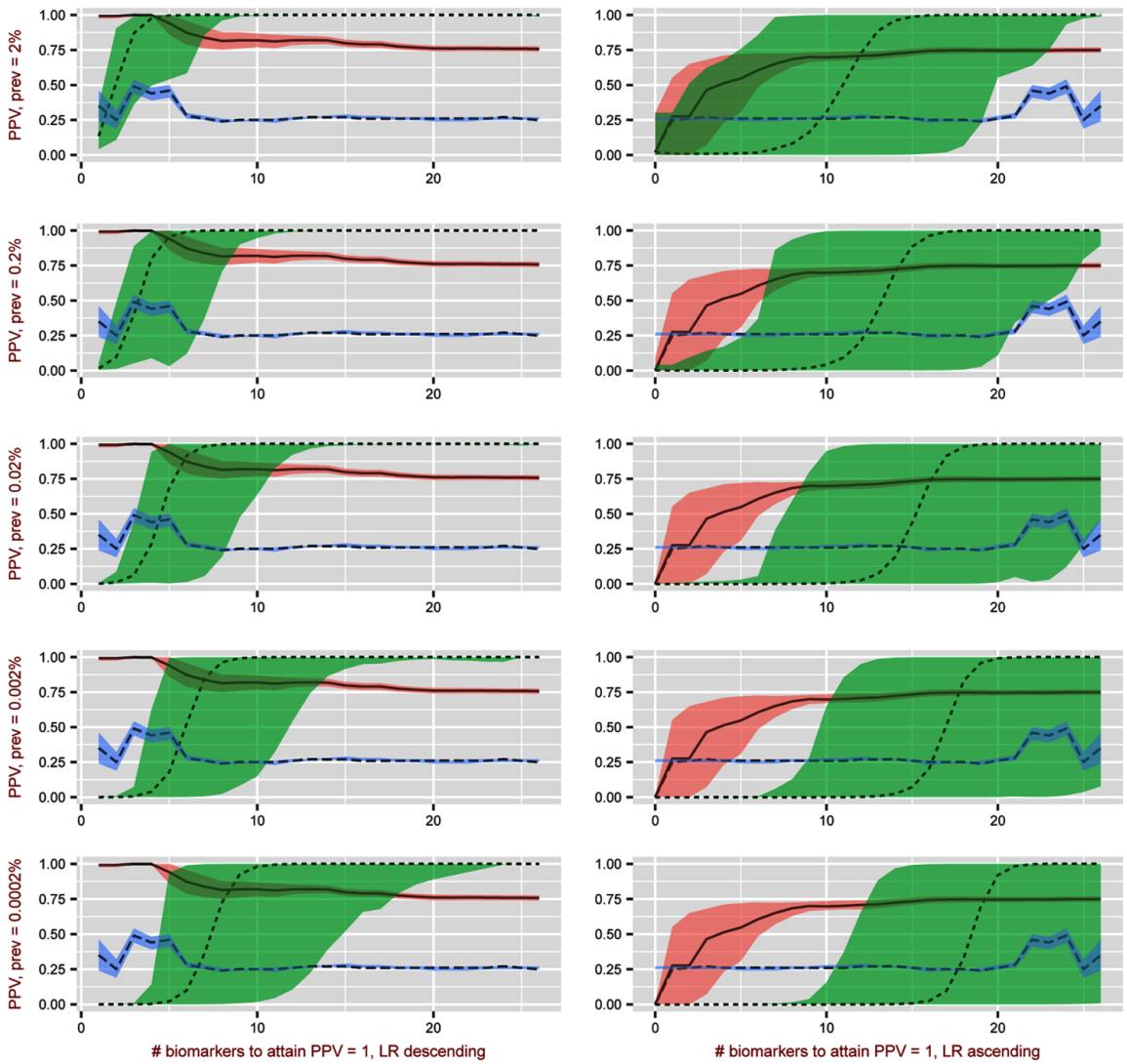

Figure 3. Number of biomarkers needed to approximate a positive predictive value of 1.00. Rows represent example prevalences with 2% at the top, decreased by a factor of 0.1 to 0.0002% prevalence at the bottom row.

ES results

Six biomarkers showed strong ES: ESR (–3.694), AIM1 (–3.361), CALCA (–3.191), HOXA9 (–3.412), RARB (–3.218), and RASSF1A (–2.873). Another 11 showed moderate evidence: IL8 (–2.333), IL1B (–2.330), DCC (–2.245), SAT (–2.171), MINT31 (–2.009), CCNA1 (–1.956), p16 (–1.892), DAPK (–1.880), TGFBR2 (–1.745), MINT1 (–1.562), and MGMT (–1.534). Seven biomarkers only were supported by only weak evidence: HIC1 (–1.410), TIMP3 (–1.375), RBM6 (–1.256), PGP9.5 (–1.218), S100A2 (–1.123), CDH1 (–1.055), and CCND2 (–1.031). The 2 biomarkers lacking evidential support were miR31 and RIZ1.

Plots in Figure 4 show the individual biomarkers and their associated confidence intervals by sensitivity and specificity with color coding for ES. Plot points were color coded according to ES: green = strong, gold = moderate, red = weak, and black = lacks evidence. The blue dots near the bottom right represent biomarkers with sensitivity over 0.75 and specificity over 0.75. The pink dots at lower right plots biomarkers with sensitivities over 0.75 or specificities over 0.75.

Based on the summary variable estimates of n = 26, ES correlated moderately with LR, (Pearson r = .66) and PPV (r = .67). Area under the curve correlated with ES at (r = .52), LR (r = .29), and PPV (r = .42). Prevalence is not associated with ES but mildly with LR (–0.35).

Discussion

Detecting cancer early has life - and cost-saving advantages over treating advanced stages. Candidate cancer biomarkers are identified by differences in biomarker concentrations between healthy individuals and patients with OSCC. Four commonly derived indices of biomarkers—sensitivity, specificity, statistical significance, and AUC—do not yield a personal risk of disease—a key feature in the future of personalized medicine. To refine screening results to derive personal probability of disease, PPVs are the metric of choice.

Retrospective case-control methods artificially inflate prevalence and thus make it easier to overestimate efficiency in biomarker disease detection by inflating PPV estimates. LRs should be adjusted for exagerated high prevalences before estimating biomarker accuracy at lower disease prevalences.

In the final step after adjusting for prevalence, PPVs of biomarkers can be combined with other biomarkers in panels to be more informative with fewer biomarkers. Passing along positive test results from one biomarker to the next refines and augments information as to true disease status of a patient.

In all, 24 of 26 studied biomarkers demonstrated at least weak evidence for OSCC detection. The other 2 biomarkers indicated that their inclusion reduced PPV, ie, increased screening uncertainty. When weak and low information value biomarkers are eliminated the number required to screen will decrease, improving efficiency and affordability.

Area under the curve was an inferior metric as compared to LR in for achieving an acceptable PPV with fewer biomarkers. In Figure 3, when using the highest LR biomarkers first, AUC accuracy tops Bayesian accumulation and Youden index. Few biomarker's high performance measured by AUC is expected, since high AUC means high sensitivity and specificity. Such ideal biomarkers may act as initial screens for disease, but may not identify the disease for which the screening is intended. This finding illustrates Kraemer's assertion that more tests about a condition does not improve diagnostic and screening accuracy.[14]

Value of biomarker information

Biomarker LR is a convenient expression of information value that can be manipulated mathematically to optimize biomarker usage. The number of biomarkers required to screen for a disease is a function of LR along with an estimated disease prevalence. When testing for life-threatening disease, the risk preference is to accept more false positives at the expense of false negatives.

The probability of having the disease given positive test results, PPV, is the ultimate objective of good medical tests. Biomarkers with both the highest sensitivity and specificity are the most accurate ones.[14] Biomarkers are more informative when included in panels including other biomarkers with high LR that incrementally increase true positives and eliminate true negatives from the screening pool.[16] High biomarker specificity appeared as more efficient than high sensitivity for screening.

Both AUC and Youden's index are hard to interpret and do not improve accuracy with additional biomarkers (Figure 3).

When screening a population where disease prevalence is unknown, prevalence might be estimated from groups with similar genetic predispositions, behavioral and dietary risks, cultural risks such as chewing betel nut, tobacco use, alcohol consumption, and other health habits and beliefs.

Interestingly, the most informative biomarkers (ones with the highest LR) turned out to have the highest specificities. Mathematically, this can be explained as the denominator of the LR—the proportion of false positives, or 1–specificity. Large specificities reduce the denominator and increase the overall LR.

The method of ES, where diagnostic accuracy with a biomarker is compared to accuracy without it, appears to ignore some promising biomarkers shown in Figure 4. CCND2, HIC1, PGP9.5, TGFBR2, and TIMP3 have more consistent grouping of data points and therefore are more reliable estimates of their AUC. However, since none of their regression lines pull towards the upper left-hand corner of high accuracy, such biomarkers are reliable but not accurate. They may be used alone or in combinations with others to achieve the best accuracy per unit costs of deployment, preservation, and transportation.

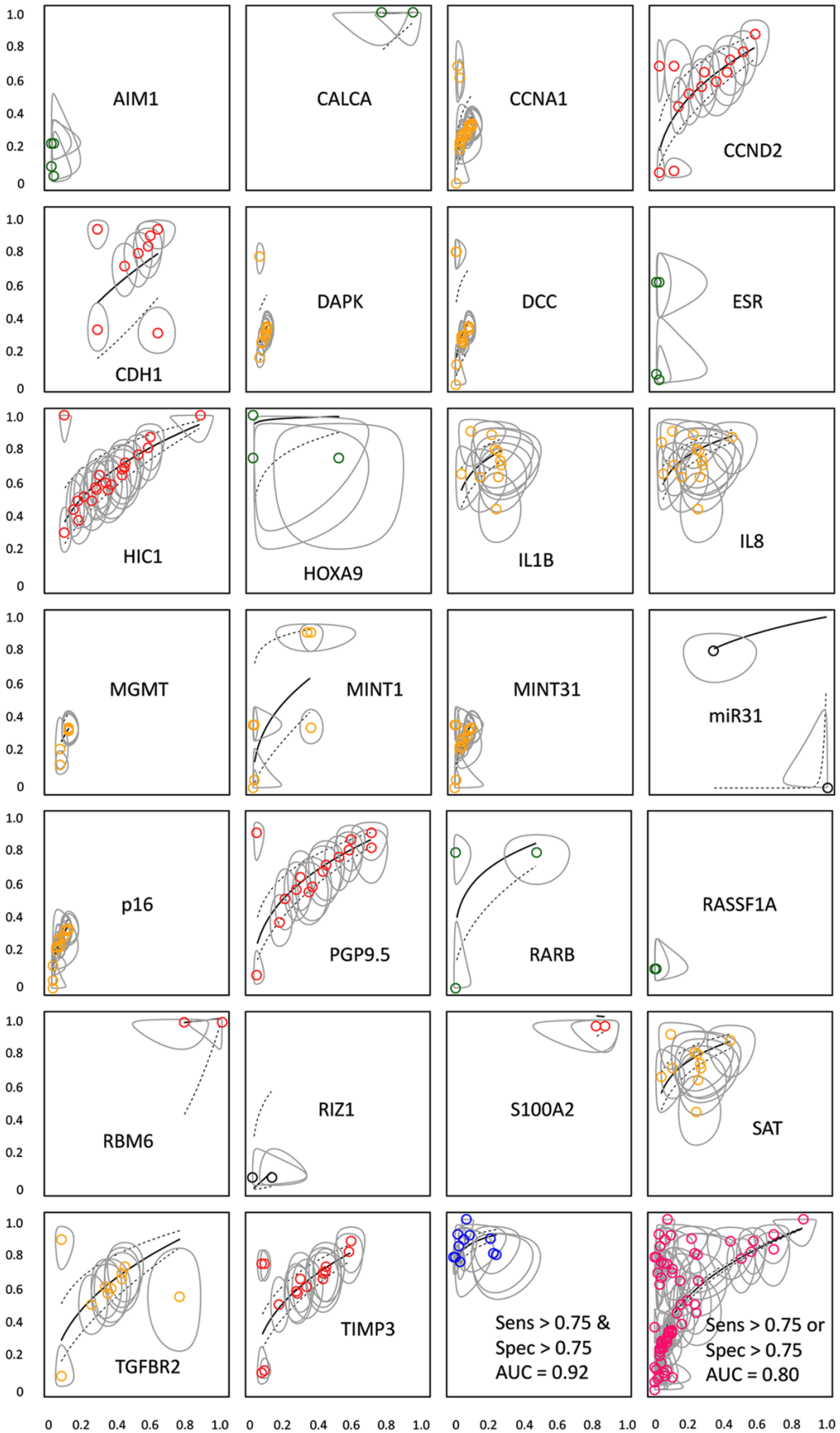

Figure 4. Individual biomarker AUC plots coded by color for evidential support: green = strong evidence, gold = moderate, red = weak, black = lack of evidence Panels in bottom right illustrate 0.75 sensitivity *and* 0.75 specificity biomarkers (blue) and 0.75 sensitivity *or* 0.75 specificity biomarkers (pink) with their respective AUC's.

Limitations of method

The estimated LR of biomarkers in panels was probably diluted by other biomarkers also in the panel. If so, then we may expect biomarker accuracy to be lower than that estimated for single biomarkers in isolation. Nonetheless, biomarkers may interact synergistically in their ability to detect disease in which cases the simultaneous presence of such biomarkers has a higher LR than that expected from each individual one.

Where physicians, patients, and public health officials have budgetary constraints for biomarkers in field settings, the number of biomarkers required for screening may be reduced by obtaining a more focused at-risk population by first screening for behaviors and family history. The risk of making errors as false negatives versus false positives must always be balanced.

Data used herein were limited to studies published through 2014. Data collection ended in 2014 so that the method could be developed without introducing new information. Information published since then may update findings and be useful to test the method.

Some biomarkers tested are conceptually too general and will not serve to identify a specific disease. The ESR is one of those biomarkers. Elevated ESR might serve as a first-level general biomarker screen to be passed along to other biomarkers for more information. Elevated ESR indicates many diseases, not just cancer. Thus, ESR's strong ES and high sensitivity may make it ideal for ruling out cancer with a negative ESR result.

Finally, assays that came from the same authors or same labs must be considered as sources of bias. One paper[63] reported all the CCND2 biomarker results, and perhaps a single laboratory was responsible for all CCND2 analyses. If so, a bias must be ruled out in future studies.

Conclusions

The same diagnostic and screening information may be obtained from fewer but more informative biomarkers rather than many weak to moderately informative ones.

By converting summary salivary biomarker data to positive likelihood ratios and then reading them as conditional probabilities ordered from highest to lowest LR, higher efficiency of OSCC screening may be obtained with fewer biomarkers. Sensitivity, specificity, and AUC do not allow the aggregation and ordering needed to improve screening accuracy.

The future and promise of personalized medicine is exactly the problem tackled here—how to translate multiple indicators of population-level disease and personal characteristics into a particular patient's probability of disease? Individual patient test results can then inform treatment and prevention strategies in the context of culture, behaviors, genetic predispositions, and economic and caretaker burdens.

REFERENCE

1. Ziliak ST, McCloskey DN. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. Ann Arbor, MI: The University of Michigan; 2011.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127(12):2893–2917.
3. Shenoi R, Devrukhkar V, Chaudhuri Sharma BK, Sapre SP, Chikhale A. Demographic and clinical profile of oral squamous cell carcinoma patients: a retrospective study. *Indian J Cancer*. 2012;49(1):21–26.
4. Zhang H, Dziegielewski PT, Biron VL, et al. Survival outcomes of patients with advanced oral cavity squamous cell carcinoma treated with multimodal therapy: a multi-institutional analysis. *J Otolaryngol Head Neck Surg*. 2013; 42:30.
5. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol*. 2009;45(4-5):309–316.
6. Farah CS, Jessri M, Currie S, Alnuaimi A, Yap T, McCullough MJ. Aetiology of oral cavity cancer. In: Kuriakose MA, ed. *Contemporary Oral Oncology*. Berlin, Germany: Springer; 2016:31–76.
7. Kruse AL, Bredell M, Gratz KW. Oral squamous cell carcinoma in non-smoking and nondrinking patients. *Head Neck Oncol*. 2010;2:24.
8. Tzioufas AG, Kapsogeorgou EK, Biomarkers: saliva proteomics is a promising tool to study Sjogren syndrome. *Nat Rev Rheumatol*. 2015;11:202–203.
9. Wu JY, Yi C, Chung HR, et al. Potential biomarkers in saliva for oral squamous cell carcinoma. *Oral Oncol*. 2010;46(4):226–231.
10. Yoshizawa JM, Schafer CA, Schafer JJ, Farrell JJ, Paster BJ, Wong DT. Salivary biomarkers: toward future clinical and diagnostic utilities. *Clin Microbiol Rev*. 2013;26(4):781–791.
11. Zhang L, Xiao H, Karlan S, et al. Discovery and preclinical validation of salivary transcriptomic and proteomic biomarkers for the non-invasive detection of breast cancer. *PLoS ONE*. 2010;5(12).
12. Brinkmann O, Wong DT. Salivary transcriptome biomarkers in oral squamous cell cancer detection. *Adv Clin Chem*. 2011;55:21–34.
13. Ai J, Smith B, Wong DT. Saliva Ontology: An ontology-based framework for a Salivaomics Knowledge Base. *BMC Bioinformatics*. 2010;11:302.
14. Kraemer HC. *Evaluating Medical Tests: Objective and Quantitative Guidelines*. 1st ed. Newbury Park, CA: SAGE; 1992.
15. Goodman SN, Royall R. Evidence and scientific research. *Am J Public Health*. 1988;78(12):1568–1574.
16. Haynes RB, Sackett DL, Guyatt GH, Tugwell P. *Clinical Epidemiology: How to Do Clinical Practice Research*. Philadelphia, PA: Lippincott Williams & Wilkins; 2006.
17. Louis TA. Introduction to Bayesian methods II: fundamental concepts. *Clin Trials*. 2005;2(4):291–294; discussion 301–294, 364–278.
18. Biggerstaff BJ. Comparing diagnostic tests: a simple graphic using likelihood ratios. *Stat Med*. 2000;19(5):649–663.
19. Pepe MS, Etzioni R, Feng ZD, et al. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst*. 2001;93(14):1054–1061.
20. Cai TX, Pepe MS. Semiparametric receiver operating characteristic analysis to evaluate biomarkers for disease. *J Am Stat Assoc*. 2002;97(460):1099–1107.
21. Dodd LE, Pepe MS. Partial AUC estimation and regression. *Biometrics*. 2003;59(3):614–623.
22. Li Y, St John MA, Zhou X, et al. Salivary transcriptome diagnostics for oral cancer detection. *Clin Cancer Res*. 2004;10(24):8442–8450.
23. St John MA, Li Y, Zhou X, et al. Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. *Arch Otolaryngol Head Neck Surg*. 2004;130(8):929–935.
24. Xiao S, Pepe MS. Evaluating markers for selecting a patient's treatment. *Biometrics*. 2004;60(4):874–883.
25. Pepe MS. Evaluating technologies for classification and prediction in medicine. *Stat Med*. 2005;24(24):3687–3696.
26. Cai TX, Pepe MS, Zheng YY, Lumley T, Jenny NS. The sensitivity and specificity of markers for event times. *Biostatistics*. 2006;7(2):182–197.
27. Hu S, Loo JA, Wong DT. Human body fluid proteome analysis. *Proteomics*. 2006;6(23):6326–6353.
28. Pepe MS, Cai TX, Longton G. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics*. 2006;62(1):221–229.
29. Hu S, Loo JA, Wong DT. Human saliva proteome analysis and disease biomarker discovery. *Expert Rev Proteomics*. 2007;4(4):531–538.
30. Hu S, Wang J, Meijer J, et al. Salivary proteomic and genomic biomarkers for primary Sjogren's syndrome. *Arthritis Rheum*. 2007;56(11):3588–3600.
31. Hu S, Yen Y, Ann D, Wong DT. Implications of salivary proteomics in drug discovery and development: a focus on cancer drug discovery. *Drug Discovery Today*. 2007;12(21-22):911–916.
32. Huang Y, Pepe MS, Feng Z. Evaluating the predictiveness of a continuous marker. *Biometrics*. 2007;63(4):1181–1188.
33. Carvalho AL, Jeronimo C, Kim MM, et al. Evaluation of promoter hypermethylation detection in body fluids as a screening/diagnosis tool for head and neck squamous cell carcinoma. *Clin Cancer Res*. 2008;14(1):97–107.
34. Hu S, Arellano M, Boontheung P, et al. Salivary proteomics for oral cancer biomarker discovery. *Clin Cancer Res*. 2008;14:6246–6252.

35. Janes H, Pepe MS. Matching in studies of classification accuracy: implications for analysis, efficiency, and assessment of incremental value. *Biometrics*. 2008;64(1):1–9.
36. Pepe MS, Feng Z, Huang Y, et al. Integrating the predictiveness of a marker with its performance as a classifier. *Am J Epidemiol*. 2008;167(3):362–368.
37. Viet CT, Schmidt BL. Methylation array analysis of preoperative and postoperative saliva DNA in oral cancer patients. *Cancer Epidemiol Biomarkers Prev*. 2008;17(12):3603–3611.
38. Gao K, Zhou H, Zhang L, et al. Systemic disease-induced salivary biomarker profiles in mouse models of melanoma and non-small cell lung cancer. *PLoS ONE*. 2009;4(6).
39. Huang Y, Pepe MS. A parametric ROC model-based approach for evaluating the predictiveness of continuous markers in case-control studies. *Biometrics*. 2009;65(4):1133–1144.
40. Huang Y, Pepe MS. Semiparametric methods for evaluating risk prediction markers in casecontrol studies. *Biometrika*. 2009;96(4):991–997.
41. Huang Y, Pepe MS. Biomarker evaluation and comparison using the controls as a reference population. *Biostatistics*. 2009;10(2):228–244.
42. Janes H, Pepe MS. Adjusting for covariate effects on classification accuracy using the covariate-adjusted receiver operating characteristic curve. *Biometrika*. 2009;96(2):371–382.
43. Lee Y-H, Wong DT. Saliva: an emerging biofluid for early detection of diseases. *Am J Dent*. 2009;22(4):241–248.
44. Park NJ, Zhou H, Elashoff D, et al. Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. *Clin Cancer Res*. 2009;15(17):5473–5477.
45. Pepe MS, Feng Z, Longton G, Koopmeiners J. Conditional estimation of sensitivity and specificity from a phase 2 biomarker study allowing early termination for futility. *Stat Med*. 2009;28(5):762–779.
46. Shaw PA, Pepe MS, Alonzo TA, Etzioni R. Methods for assessing improvement in specificity when a biomarker is combined with a standard screening test. *Stat Biopharm Res*. 2009;1(1):18–25.
47. Wei F, Patel P, Liao W, et al. Electrochemical sensor for multiplex biomarkers detection. *Clin Cancer Res*. 2009;15(13):4446–4452.
48. Yan W, Apweiler R, Balgley BM, et al. Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin Appl*. 2009;3(1):116–134.
49. Zhang L, Xiao H, Wong DT. Salivary biomarkers for clinical applications. *Mol Diagn Ther*. 2009;13(4):245–259.
50. Demokan S, Chang X, Chuang A, et al. KIF1A and EDNRB are differentially methylated in primary HNSCC and salivary rinses. *Int J Cancer*. 2010;127(10):2351–2359.
51. Halgand F, Zabrouskov V, Bassilian S, et al. Micro-heterogeneity of human saliva peptide P-C characterized by high-resolution top-down Fourier-transform mass spectrometry. *J Am Soc Mass Spectrom*. 2010;21(5):868–877.
52. Hu S, Jiang J, Wong DT. Proteomic analysis of saliva: 2D gel electrophoresis, LC-MS/MS, and western blotting. In: Seymour GJ, Cullinan MP, Heng NCK, eds. *Oral Biology: Molecular Techniques and Applications*. Vol 666. Berlin, Germany: Springer; 2010:31–41.
53. Huang Y, Pepe MS. Assessing risk prediction models in case-control studies using semiparametric and nonparametric methods. *Stat Med*. 2010;29(13):1391–1410.
54. Huang Y, Pepe MS. Semiparametric methods for evaluating the covariate-specific predictiveness of continuous markers in matched case-control studies. *J R Stat Soc Ser C Appl Stat*. 2010;59:437–456.
55. Kruse AL, Bredell M, Luebbers HT, Gratz KW. Head and neck cancer in the elderly: a retrospective study over 10 years (1999-2008). *Head Neck Oncol*. 2010;2:25.
56. Morris DE, Pepe MS, Barlow WE. Contrasting two frameworks for ROC analysis of ordinal ratings. *Med Decis Making*. 2010;30(4):484–498.
57. Palanisamy V, Wong DT. Transcriptomic analyses of saliva. In: Seymour GJ, Cullinan MP, Heng NCK, eds. *Oral Biology: Molecular Techniques and Applications*. Vol 666. Berlin, Germany: Springer; 2010:43–51.
58. Pattani KM, Zhang Z, Demokan S, et al. Endothelin receptor type B gene promoter hypermethylation in salivary rinses is independently associated with risk of oral cavity cancer and premalignancy. *Cancer Prev Res (Phila)*. 2010;3(9):1093–1103.
59. Ryan CM, Souda P, Halgand F, et al. Confident assignment of intact mass tags to human salivary cystatins using top-down Fourier-transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom*. 2010;21(6):908–917.
60. Sugimoto M, Wong DT, Hirayama A, Soga T, Tomita M. Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancerspecific profiles. *Metabolomics*. 2010;6(1):78–95.
61. Zhang L, Farrell JJ, Zhou H, et al. Salivary transcriptomic biomarkers for detection of resectable pancreatic cancer. *Gastroenterology*. 2010;138(3): 949–957.
62. Brinkmann O, Kastratovic DA, Dimitrijevic MV, et al. Oral squamous cell carcinoma detection by salivary biomarkers in a Serbian population. *Oral Oncology*. 2011;47(1):51–55.
63. Carvalho AL, Henrique R, Jeronimo C, et al. Detection of promoter hypermethylation in salivary rinses as a biomarker for head and neck squamous cell carcinoma surveillance. *Clin Cancer Res*. 2011;17(14):4782–4789.
64. Cheng Y-SL, Rees T, Jordan L, et al. Salivary endothelin-1 potential for detecting oral cancer in patients with oral lichen planus or oral cancer in remission. *Oral Oncology*. 2011;47(12):1122–1126.
65. Gu W, Pepe MS. Estimating the diagnostic likelihood ratio of a continuous marker. *Biostatistics*. 2011;12(1):87–101.
66. Guerrero-Preston R, Herbstman J, Goldman LR. Epigenomic biomonitors: global DNA hypomethylation as a biodosimeter of life-long environmental exposures. *Epigenomics*. 2011;3(1):1–5.
67. Guerrero-Preston R, Soudry E, Acero J, et al. NID2 and HOXA9 promoter hypermethylation as biomarkers for prevention and early detection in oral cavity squamous cell carcinoma tissues and saliva. *Cancer Prev Res (Phila)*. 2011;4(7):1061–1072.
68. Hu S, Vissink A, Arellano M, et al. Identification of autoantibody biomarkers for primary Sjogren's syndrome using protein microarrays. *Proteomics*. 2011;11(8):1499–1507.
69. Keidar M, Walk R, Shashurin A, et al. Cold plasma selectivity and the possibility of a paradigm shift in cancer therapy. *Br J Cancer*. 2011;105(9):1295–1301.
70. Kruse AL, Bredell M, Gratz KW. Oral cancer in men and women: are there differences? *Oral Maxillofac Surg*. 2011;15(1):51–55.
71. Spielmann N, Wong DT. Saliva: diagnostics and therapeutic perspectives. *Oral Dis*. 2011;17(4):345–354.
72. Wei J, Xie G, Zhou Z, et al. Salivary metabolite signatures of oral cancer and leukoplakia. *Int J Cancer*. 2011;129(9):2207–2217.
73. Bonne NJ, Wong DTW. Salivary biomarker development using genomic, proteomic and metabolomic approaches. *Genome Med*. 2012;4:82.
74. Elashoff D, Zhou H, Reiss J, et al. Prevalidation of salivary biomarkers for oral cancer detection. *Cancer Epidem Biomar*. 2012;21(4):664–672.
75. Farrell JJ, Zhang L, Zhou H, et al. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. *Gut*. 2012;61(4): 582–588.
76. Koopmeiners JS, Feng Z, Pepe MS. Conditional estimation after a two-stage diagnostic biomarker study that allows early termination for futility. *Stat Med*. 2012;31(5):420–435.
77. Lee Y-H, Kim JH, Zhou H, Kim BW, Wong DT. Salivary transcriptomic biomarkers for detection of ovarian cancer: for serous papillary adenocarcinoma. *JMol Med-JMM*.2012;90(4):427–434.
78. Liu CJ, Lin SC, Yang CC, Cheng HW, Chang KW. Exploiting salivary miR-31 as a clinical biomarker of oral squamous cell carcinoma. *Head Neck*. 2012;34(2):219–224.
79. Pepe MS, Fan J, Seymour CW, Li C, Huang Y, Feng Z. Biases introduced by choosing controls to match risk factors of cases in biomarker research. *Clin Chem*. 2012;58(8):1242–1251.
80. Xiao H, Wong DTW. Proteomic analysis of microvesicles in human saliva by gel electrophoresis with liquid chromatography-mass spectrometry. *Anal Chim Acta*. 2012;723:61–67.
81. Xiao H, Zhang L, Zhou H, Lee JM, Garon EB, Wong DTW. Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. *Mol Cell Proteomics*. 2012;11(2).
82. Zhang L, Xiao H, Zhou H, et al. Development of transcriptomic biomarker signature in human saliva to detect lung cancer. *Cell Mol Life Sci*. 2012;69(19):3341–3350.
83. Bansal A, Pepe MS. Estimating improvement in prediction with matched case-control designs. *Lifetime Data Anal*. 2013;19(2):170–201.
84. Matse JH, Yoshizawa J, Wang X, et al. Discovery and prevalidation of salivary extracellular microRNA biomarkers panel for the noninvasive detection of benign and malignant parotid gland tumors. *Clin Cancer Res*. 2013;19(11):3032–3038.
85. Pepe MS, Kerr KF, Longton G, Wang Z. Testing for improvement in prediction model performance. *Stat Med*. 2013;32(9):1467–1482.
86. Wei F, Yang J, Wong DTW. Detection of exosomal biomarker by electric field-induced release and measurement (EFIRM). *Biosens Bioelectron*. 2013;44:115–121.
87. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31.
88. *R: A Language and Environment for Statistical Computing* (computer program). Vienna,Austria: R Foundation for Statistical Computing; 2017.
89. Menke JM. Do manual therapies help low back pain? A comparative effectiveness metaanalysis. *Spine (Phila PA 1976)*. 2014;39(7):E463–472.
90. Bryk AS, Raudenbush SW. *Hierarchical Linear Models*. Newbury Park, CA: SAGE; 1992.
91. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58(10):982–990.
92. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making*. 1993; 1993(13):313–321.

Appendix 1

All biomarker panels culled from 16 evidence summary papers.

BIOMARKER PANEL	REFERENCE NUMBER	NUMBER OF PATIENTS WITH OSCC	NUMBER OF PATIENTS WITHOUT OSCC	TOTAL PATIENTS	TP	FP	FN	TN	SENSITIVITY	SPECIFICITY	LR+
ADCYAP1_P455_R + CEBPA_P706_F + EPHA5_E158_R + FGF3_ E198_R + HLF_ E192_F + IL11_P11_R + INSR_P1063_R + NOTCH3_ E403_F	37	13	10	23	9	0	4	10	0.690 (0.562 to 0.818)	0.960 (0.898 to 1.00)	17.250
AIM1	63	23	73	96	1	1	22	72	0.044 (0.001 to 0.220)	0.986 (0.926 to 1.00)	3.143
AIM1	63	10	41	51	1	0	9	41	0.100 (0.003 to 0.445)	0.990 (0.930 to 1.00)	10.000
AIM1	63	77	41	118	18	0	59	41	0.234 (0.145 to 0.344)	0.990 (0.930 to 1.00)	23.400
AIM1	63	77	73	150	18	1	59	72	0.234 (0.145 to 0.344)	0.986 (0.926 to 1.00)	16.714
CALCA	63	35	20	55	35	15	0	5	1.000 (0.918 to 1.000)	0.250 (0.087 to 0.491)	1.333
CALCA	63	35	30	65	35	28	0	2	1.000 (0.918 to 1.000)	0.067 (0.008 to 0.221)	1.072
CCNA1	63	24	104	128	0	0	24	104	0.000 (0.000 to 0.117)	0.990 (0.972 to 1.00)	0.000
CCNA1	63	175	444	619	35	13	140	431	0.200 (0.143 to 0.267)	0.971 (0.951 to 0.984)	6.897
CCNA1	63	102	444	546	63	13	39	431	0.618 (0.516 to 0.712)	0.971 (0.951 to 0.984)	21.310
CCNA1	63	35	284	319	24	5	11	279	0.686 (0.516 to 0.712)	0.982 (0.972 to 1.00)	38.111
CCNA1 + DCC + DAPK + MINT31 + p16	33	211	527	738	72	43	139	484	0.341 (0.308 to 0.374)	0.918 (0.630 to 1.00)	4.159
CCNA1 + DAPK	63	176	416	592	49	30	127	386	0.278 (0.214 to 0.351)	0.928 (0.889 to 0.951)	3.861
CCNA1 + DAPK + p16	63	176	416	592	52	30	124	386	0.295 (0.229 tp 0.369)	0.928 (0.889 to 0.951)	4.097
CCNA1 + DCC	63	175	444	619	45	18	130	426	0.257 (0.194 to 0.329)	0.959 (0.937 to 0.976)	6.268
CCNA1 + DCC + DAPK	63	176	417	593	56	34	120	383	0.318 (0.250 to 0.393)	0.918 (0.888 to 0.943)	3.878
CCNA1 + DCC + DAPK + p16	63	176	417	593	59	34	117	383	0.335 (0.266 to 0.410)	0.918 (0.888 to 0.943)	4.085
CCNA1 + DCC + p16	63	175	444	619	49	18	126	426	0.280 (0.215 to 0.353)	0.959 (0.937 to 0.976)	6.829
CCNA1 + p16	63	175	444	619	39	13	136	431	0.223 (0.164 to 0.292)	0.971 (0.951 to 0.984)	7.690
CCND2	63	47	284	331	3	5	44	279	0.064 (0.013 to 0.175)	0.982 (0.959 to 0.994)	3.556
CCND2	63	136	97	233	10	10	126	87	0.074 (0.036 to 0.131)	0.897 (0.819 to 0.949)	0.718
CCND2	63	35	97	132	24	10	11	87	0.686 (0.507 to 0.832)	0.897 (0.819 to 0.947)	6.660
CCND2	63	35	284	319	24	5	11	279	0.686 (0.507 to 0.832)	0.982 (0.959 to 0.994)	38.111
CCND2 + HIC1	63	49	248	297	22	32	27	216	0.449 (0.307 to 0.598)	0.871 (0.823 to 0.910)	3.481
CCND2 + HIC1 + PGP9.5	63	42	189	231	22	36	20	153	0.524 (0.364 to 0.680)	0.810 (0.746 to 0.863)	2.758

(Continued)

(Continued)

BIOMARKER PANEL	REFERENCE NUMBER	NUMBER OF PATIENTS WITH OSCC	NUMBER OF PATIENTS WITHOUT OSCC	TOTAL PATIENTS	TP	FP	FN	TN	SENSITIVITY	SPECIFICITY	LR+
CCND2 + TGFBR2 + TIMP3 + HIC1 + PGP9.5	63	36	130	166	26	56	10	74	0.722 (0.548 to 0.858)	0.569 (0.480 to 0.656)	1.675
CCND2 + TGFBR2 +TIMP3 + HIC1	63	40	126	166	26	52	14	74	0.650 (0.483 to 0.794)	0.587 (0.496 to 0.674)	1.574
CCND2 + TGFRB2 + HIC1 + PGP9.5	63	37	117	154	22	41	15	76	0.595 (0.421 to 0.753)	0.650 (0.556 to 0.736)	1.700
CCND2 + TIMP3 + HIC1	63	46	178	224	26	47	20	131	0.565 (0.411 to 0.711)	0.736 (0.665 to 0.799)	2.140
CCND2 + TIMP3 + HIC1 + PGP9.5	63	40	182	222	26	51	14	131	0.650 (0.483 to 0.794)	0.720 (0.649 to 0.784)	2.321
CDH1	63	66	116	182	20	72	46	44	0.303 (0.196 to 0.429)	0.379 (0.291 to 0.474)	0.488
CDH1	63	62	320	382	20	86	42	234	0.323 (0.209 to 0.453)	0.731 (0.679 to 0.779)	1.201
CDH1	63	77	116	193	70	72	7	44	0.909 (0.822 to 0.963)	0.379 (0.291 to 0.474)	1.464
CDH1	63	77	320	397	70	86	7	234	0.909 (0.822 to 0.963)	0.731 (0.679 to 0.779)	3.379
CDH1 + CCND2 + HIC1 + PGP9.5	63	39	217	256	30	110	9	107	0.769 (0.607 to 0.889)	0.493 (0.425 to 0.562)	1.517
CDH1 + CCND2 + TIMP3 + HIC1 + PGP9.5	63	39	208	247	34	120	5	88	0.872 (0.726 to 0.957)	0.423 (0.355 to 0.493)	1.511
CDH1 + TIMP3 + HIC1	63	49	267	316	34	114	15	153	0.694 (0.546 to 0.818)	0.573 (0.511 to 0.633)	1.625
CDH1 + TIMP3 + HIC1 + PGP9.5	63	42	209	251	34	118	8	91	0.810 (0.659 to 0.914)	0.435 (0.367 to 0.506)	1.434
DAPK	63	176	451	627	28	17	148	434	0.159 (0.108 to 0.222)	0.962 (0.940 to 0.978)	4.184
DAPK	63	136	451	587	102	17	34	434	0.750 (0.669 to 0.820)	0.962 (0.940 to 0.978)	19.737
DCC	63	27	135	162	0	0	27	135	0.000 (0.000 to 0.105)	0.990 (0.978 to 1.00)	0.000
DCC	63	176	462	638	21	5	155	457	0.119 (0.075 to 0.177)	0.989 (0.975 to 0.997)	10.818
DCC	63	135	135	270	105	0	30	135	0.778 (0.698 to 0.845)	0.990 (0.978 to 1.00)	77.800
DCC	63	135	462	597	105	5	30	457	0.778 (0.698 to 0.845)	0.989 (0.975 to 0.997)	70.727
DCC + DAPK + p16	63	176	422	598	43	21	133	401	0.244 (0.183 to 0.315)	0.950 (0.925 to 0.969)	4.880
EDNRB	58	161	30	191	105	15	56	15	0.650 (0.612 to 0.688)	0.510 (0.419 to 0.601)	1.327
ERBB4_P255_F + IL11_P11_R + PTCH2_ P37_F + TMEFF1_P234_F + TNFSF10_E53_F + TWIST1_P44_R	37	13	10	23	8	0	5	10	0.620 (0.485 to 0.755)	0.990 (0.959 to 1.00)	62.000
ESR	63	33	119	152	1	2	32	117	0.030 (0.001 to 0.158)	0.983 (0.941 to 0.998)	1.765
ESR	63	16	20	36	1	0	15	20	0.063 (0.002 to 0.302)	0.990 (0.861 to 1.00)	6.300
ESR	63	35	20	55	21	0	14	20	0.600 (0.421 to 0.761)	0.990 (0.861 to 1.00)	60.000
ESR	63	35	119	154	21	2	14	117	0.600 (0.421 to 0.761)	0.983 (0.941 to 0.998)	35.294
GABRB3_E42_F + IL11_P11_R + INSR_ P1063_R + NOTCH3_E403_F + NTRK3_ E131_F + PXN_P308_F	37	13	10	23	10	2	3	8	0.770 (0.653 to 0.887)	0.830 (0.711 to 0.949)	4.529

(Continued)

BIOMARKER PANEL	REFERENCE NUMBER	NUMBER OF PATIENTS WITH OSCC	NUMBER OF PATIENTS WITHOUT OSCC	TOTAL PATIENTS	TP	FP	FN	TN	SENSITIVITY	SPECIFICITY	LR+
HIC1	63	70	373	443	22	28	48	345	0.314 (0.209 to 0.436)	0.925 (0.893 to 0.950)	4.187
HIC1	63	45	46	91	45	40	0	6	1.000 (0.936 to 1.000)	0.130 (0.049 to 0.263)	1.149
HIC1	63	45	373	418	45	28	0	345	1.000 (0.936 to 1.000)	0.925 (0.893 to 0.950)	13.333
HIC1 + PGP9.5	63	57	202	259	22	32	35	170	0.386 (0.260 to 0.524)	0.842 (0.784 to 0.889)	2.443
HOXA9	67	4	4	8	3	0	1	4	0.680 (0.447 to 0.913)	0.990 (0.940 to 1.00)	68.000
HOXA9 + NID2	67	4	4	8	4	0	0	4	0.940 (0.821 to 1.00)	0.970 (0.885 to 1.00)	31.333
IL-1B + IL-8 + M2BP	62	35	51	86	31	11	4	40	0.890 (0.837 to 0.943)	0.780 (0.722 to 0.838)	4.045
IL-8	23	19	32	51	16	1	3	31	0.860 (0.780 to 0.940)	0.970 (0.940 to 1.00)	28.667
IL-8 + IL-1B + SAT + S100P	6,8	35	51	86	24	2	12	49	0.670 (0.591 to 0.749)	0.960 (0.705 to 1.00)	16.750
IL8 + IL1B + OAZ1 + SAT	22	32	32	64	29	3	3	29	0.910 (0.859 to 0.961)	0.910 (0.601 to 1.00)	10.111
IL8 + IL1B + SAT + DUSP1	74 [cohort 3]	30	30	60	24	7	6	23	0.800 (0.727 to 0.873)	0.770 (0.276 to 1.00)	3.478
IL8 + IL1B + SAT + OAZ1	74 [cohort 5]	31	70	101	14	17	17	53	0.450 (0.361 to 0.539)	0.760 (0.049 tp 1.00)	1.875
IL8 + IL1B + SAT + OAZ1	74 [cohort 4]	36	54	90	23	14	13	41	0.640 (0.560 to 0.720)	0.750 (0.171 to 1.00)	2.560
IL8 + IL1B + SAT + OAZ1	74 [cohort 1]	48	48	96	34	13	14	35	0.700 (0.634 to 0.766)	0.720 (0.156 to 1.00)	2.500
IL8 + IL1B + SAT + OAZ1	74 [cohort 3]	30	30	60	22	8	8	22	0.730 (0.649 to 0.811)	0.730 (0.179 to 1.00)	2.704
IL8 + IL1B + SAT + OAZ1	74 [cohort 2]	24	24	48	19	6	5	19	0.790 (0.707 to 0.873)	0.770 (0.269 to 1.00)	3.435
IL8 + SAT + H3F3A + S100P	74 [cohort 1]	48	48	96	34	5	14	43	0.710 (0.645 to 0.775)	0.890 (0.500 to 1.00)	6.455
IL8 + SAT + OAZ1 + S100P	74 [cohort 5]	31	70	101	27	31	4	39	0.870 (0.810 to 0.930)	0.560 (0.008 to 1.00)	1.977
IL8+IL1B+S100P+OAZ1	74 [cohort 4]	36	54	90	23	8	13	46	0.640 (0.560 to 0.720)	0.860 (0.813 to 0.907)	4.571
M2BP + profilin + CD59 + MRP14 + catalase	34	64	64	128	58	11	6	53	0.900 (0.863 to 0.938)	0.830 (0.783 to 0.877)	5.294
MGMT	63	149	239	388	20	12	129	227	0.134 (0.084 to 0.200)	0.950 (0.914 to 0.974)	2.680
MGMT	63	44	239	283	10	12	34	227	0.227 (0.115 to 0.378)	0.950 (0.914 to 0.974)	4.540
MGMT + CCNA1	63	150	240	390	50	24	100	216	0.333 (0.259 to 0.415)	0.900 (0.855 to 0.935)	3.330
MGMT + CCNA1 + p16	63	151	240	391	51	24	100	216	0.338 (0.263 to 0.419)	0.900 (0.855 to 0.935)	3.380
MINT1	63	131	296	427	46	100	85	196	0.351 (0.270 to 0.439)	0.662 (0.605 to 0.716)	1.038
MINT1	63	87	19	106	79	6	8	13	0.908 (0.827 to 0.960)	0.684 (0.435 to 0.874)	2.873
MINT1	63	87	296	383	79	100	8	196	0.908 (0.827 to 0.960)	0.662 (0.605 to 0.716)	2.686

(Continued)

(Continued)

BIOMARKER PANEL	REFERENCE NUMBER	NUMBER OF PATIENTS WITH OSCC	NUMBER OF PATIENTS WITHOUT OSCC	TOTAL PATIENTS	TP	FP	FN	TN	SENSITIVITY	SPECIFICITY	LR+
MINT31	63	28	42	70	0	0	28	42	0.000 (0.000 to 0.102)	0.990 (0.931 to 1.00)	0.000
MINT31	63	175	492	667	8	0	167	492	0.046 (0.020 to 0.088)	0.990 (0.994 to 1.00)	4.600
MINT31	63	136	42	178	50	0	86	42	0.368 (0.287 to 0.455)	0.990 (0.931 to 1.00)	36.800
MINT31	63	136	492	628	50	0	86	492	0.368 (0.287 to 0.455)	0.990 (0.994 to 1.00)	36.800
MINT31 + CCNA1	63	174	444	618	39	13	135	431	0.224 (0.165 to 0.293)	0.971 (0.951 to 0.984)	7.724
MINT31 + CCNA1 + DAPK	63	175	416	591	52	30	123	386	0.297 (0.231 to 0.371)	0.928 (0.889 to 0.951)	4.125
MINT31 + CCNA1 + DAPK + p16	63	176	416	592	54	30	122	386	0.307 (0.240 to 0.381)	0.928 (0.889 to 0.951)	4.264
MINT31 + CCNA1 + DCC	63	174	444	618	47	18	127	426	0.270 (0.206 to 0.343)	0.959 (0.937 to 0.976)	6.585
MINT31 + CCNA1 + DCC + DAPK	63	175	417	592	58	34	117	383	0.331 (0.262 to 0.406)	0.918 (0.888 to 0.943)	4.037
MINT31 + CCNA1 + p16	63	175	444	619	42	13	133	431	0.240 (0.179 to 0.310)	0.971 (0.951 to 0.984)	8.276
MINT31 + CCNA1 + p16	50	33	61	94	8	2	25	59	0.240 (0.166 to 0.314)	0.971 (0.610 to 1.00)	8.276
MINT31 + DCC + DAPK + p16	63	176	422	598	44	21	132	401	0.250 (0.188 to 0.321)	0.950 (0.925 to 0.969)	5.000
MINT31 + MGMT + CCNA1	63	150	240	390	52	24	98	216	0.347 (0.271 to 0.429)	0.900 (0.855 to 0.935)	3.470
MINT31 + MGMT + CCNA1 + p16	63	151	240	391	53	24	98	216	0.351 (0.275 to 0.433)	0.900 (0.855 to 0.935)	3.510
MINT31 +CCNA1 + DCC + DAPK + p16	63	176	417	593	60	34	116	383	0.341 (0.271 to 0.416)	0.918 (0.888 to 0.943)	4.159
MINT31 +CCNA1 + DCC +p16	63	175	444	619	50	18	125	426	0.286 (0.220 to 0.359)	0.959 (0.937 to 0.976)	6.976
miR-125a + miR-200a	44	50	50	100	NA	NA	NA	NA	0.910 (0.870 to 0.950)	0.910 (0.870 to 0.950)	NA
miR-31	78	45	24	69	36	8	9	16	0.800 (0.740 to 0.860)	0.680 (0.585 to 0.775)	2.500
p16	63	39	102	141	0	0	39	102	0.000 (0.000 to 0.074)	0.990 (0.971 to 1.00)	0.000
p16	63	177	500	677	8	0	169	500	0.045 (0.020 to 0.087)	0.990 (0.994 to 1.00)	4.500
p16	63	136	102	238	18	0	118	102	0.132 (0.080 to 0.201)	0.990 (0.971 to 1.00)	13.200
p16	63	136	500	636	18	0	118	500	0.132 (0.080 to 0.201)	0.990 (0.994 to 1.00)	13.200
PGP9.5	63	52	203	255	4	5	48	198	0.077 (0.021 to 0.185)	0.975 (0.944 to 0.992)	3.080
PGP9.5	63	34	112	146	28	78	6	34	0.824 (0.655 to 0.932)	0.304 (0.220 to 0.398)	1.184
PGP9.5	63	45	112	157	41	78	4	34	0.911 (0.788 to 0.975)	0.304 (0.220 to 0.398)	1.309
PGP9.5	63	45	203	248	41	5	4	198	0.911 (0.788 to 0.975)	0.975 (0.944 to 0.992)	36.440
Prevalence screen HOXA9	67,66	4	4	8	3	0	1	4	0.850 (0.671 to 1.00)	0.970 (0.885 to 1.00)	28.333
RARB	63	13	85	98	0	0	13	85	0.000 (0.000 to 0.206)	0.990 (0.956 to 1.00)	0.000

(Continued)

BIOMARKER PANEL	REFERENCE NUMBER	NUMBER OF PATIENTS WITH OSCC	NUMBER OF PATIENTS WITHOUT OSCC	TOTAL PATIENTS	TP	FP	FN	TN	SENSITIVITY	SPECIFICITY	LR+
RARB	63	44	35	79	35	17	9	18	0.796 (0.647 to 0.902)	0.514 (0.340 to 0.686)	1.638
RARB	63	44	85	129	35	0	9	85	0.796 (0.647 to 0.902)	0.990 (0.965 to 1.00)	79.600
RASSF1A	63	44	35	79	5	0	39	35	0.114 (0.038 to 0.246)	0.990 (0.918 to 1.00)	11.400
RASSF1A	63	44	104	148	5	1	39	103	0.114 (0.038 to 0.246)	0.990 (0.948 to 1.00)	11.400
RBM6	63	44	18	62	43	14	1	4	0.977 (0.880 to 1.000)	0.222 (0.064 to 0.476)	1.256
RBM6	63	44	35	79	43	35	1	0	0.977 (0.880 to 1.000)	0.000 (0.000 to 0.082)	0.977
RIZ1	63	44	18	62	3	0	41	18	0.068 (0.014 to 0.187)	0.990 (0.847 to 1.00)	6.800
RIZ1	63	44	35	79	3	4	41	31	0.068 (0.014 to 0.187)	0.886 (0.733 to 0.968)	0.596
S100A2	63	44	12	56	42	10	2	2	0.955 (0.845 to 0.994)	0.167 (0.021 to 0.484)	1.146
S100A2	63	44	35	79	42	31	2	4	0.955 (0.845 to 0.994)	0.114 (0.032 to 0.267)	1.078
saliva HOXA9	67,66	4	4	8	3	2	1	2	0.750 (0.533 to 0.967)	0.530 (0.280 to 0.780)	1.596
saliva NID2	67,66	4	4	8	4	3	1	1	0.870 (0.702 to 1.00)	0.210 (0.006 to 0.414)	1.101
TGFBR2	63	37	134	171	3	8	34	126	0.081 (0.017 to 0.219)	0.940 (0.886 to 0.974)	1.350
TGFBR2	63	11	44	55	6	33	5	11	0.546 (0.234 to 0.833)	0.250 (0.132 to 0.403)	0.728
TGFBR2	63	42	134	176	37	8	5	126	0.881 (0.744 to 0.960)	0.940 (0.886 to 0.974)	14.683
TGFBR2 + HIC1	63	44	149	193	22	35	22	114	0.500 (0.346 to 0.654)	0.765 (0.689 to 0.831)	2.128
TGFBR2 + HIC1 + PGP9.5	63	39	118	157	22	39	17	79	0.564 (0.386 to 0.722)	0.669 (0.577 to 0.753)	1.704
TGFBR2 + TIMP3 + HIC1	63	43	158	201	26	50	17	108	0.605 (0.444 to 0.750)	0.684 (0.605 to 0.755)	1.915
TGFBR2 + TIMP3 + HIC1 + PGP9.5	63	38	131	169	26	54	12	77	0.684 (0.514 to 0.825)	0.588 (0.499 to 0.673)	1.660
TIMP3	63	50	296	346	5	16	45	280	0.100 (0.033 to 0.218)	0.946 (0.914 to 0.969)	1.852
TIMP3	63	176	450	626	20	32	156	418	0.114 (0.071 to 0.170)	0.929 (0.901 to 0.951)	1.606
TIMP3	63	138	296	434	102	16	36	280	0.739 (0.658 to 0.810)	0.946 (0.914 to 0.969)	13.685
TIMP3	63	138	450	588	102	32	36	418	0.739 (0.658 to 0.810)	0.929 (0.901 to 0.951)	10.408
TIMP3 + HIC1	63	52	278	330	26	43	26	235	0.500 (0.358 to 0.642)	0.845 (0.797 to 0.886)	3.226
TIMP3 + HIC1 + PGP9.5	63	45	183	228	26	47	19	136	0.578 (0.422 to 0.723)	0.743 (0.674 to 0.805)	2.249
Valine + lactic acid	72	37	32	69	32	6	5	26	0.865 (0.809 to 0.921)	0.824 (0.757 to 0.891)	4.915