VOLUME 24 · NUMBER 31 · NOVEMBER 1 2006

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Meta-Analysis and Meta-Review of Thyroid Cancer Gene Expression Profiling Studies Identifies Important Diagnostic Biomarkers

Obi L. Griffith, Adrienne Melck, Steven J.M. Jones, and Sam M. Wiseman

From the Michael Smith Genome Sciences Centre, British Columbia Cancer Agency; Departments of Medical Genetics and Surgery, University of British Columbia; Genetic Pathology Evaluation Center, Prostate Research Center of Vancouver General Hospital & British Columbia Cancer Agency; Department of Surgery, St Paul's Hospital, Vancouver, Canada.

Submitted March 27, 2006; accepted August 30, 2006.

Supported by the BC Cancer Foundation. Canadian Institutes of Health Research (O.L.G), the Natural Sciences and Engineering Council of Canada, and the Michael Smith Foundation for Health Research (S.J.M.J. and S.M.W.).

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Address reprint requests to Sam M. Wiseman, MD, FRCSC, FACS, St Paul's Hospital, Room C302 Burrard Bldg, 1081 Burrard Street, Vancouver, BC, V6Z1Y6, Canada; smwiseman@providencehealth.bc.ca.

0732-183X/06/2431-5043/$20.00

DOI: 10.1200/JCO.2006.06.7330

ABSTRACT

Purpose

An estimated 4% to 7% of the population will develop a clinically significant thyroid nodule during their lifetime. In many cases, preoperative diagnoses by needle biopsy are inconclusive. Thus, there is a clear need for improved diagnostic tests to distinguish malignant from benign thyroid tumors. The recent development of high-throughput molecular analytic techniques should allow the rapid evaluation of new diagnostic markers. However, researchers are faced with an overwhelming number of potential markers from numerous thyroid cancer expression profiling studies.

Materials and Methods

To address this challenge, we have carried out a comprehensive meta-review of thyroid cancer biomarkers from 21 published studies. A gene ranking system that considers the number of comparisons in agreement, total number of samples, average fold-change and direction of change was devised.

Results

We have observed that genes are consistently reported by multiple studies at a highly significant rate ($P < .05$). Comparison with a meta-analysis of studies reprocessed from raw data showed strong concordance with our method.

Conclusion

Our approach represents a useful method for identifying consistent gene expression markers when raw data are unavailable. A review of the top 12 candidates revealed well known thyroid cancer markers such as *MET*, *TFF3*, *SERPINA1*, *TIMP1*, *FN1*, and *TPO* as well as relatively novel or uncharacterized genes such as *TGFA*, *QPCT*, *CRABP1*, *FCGBP*, *EPS8* and *PROS1*. These candidates should help to develop a panel of markers with sufficient sensitivity and specificity for the diagnosis of thyroid tumors in a clinical setting.

J Clin Oncol 24:5043-5051. © 2006 by American Society of Clinical Oncology

INTRODUCTION

Thyroid nodules are extremely common, being palpable in 4% to 7% of the North American adult population, with new nodules detected at a yearly rate of 0.1%.[1,2] Currently, fine-needle aspiration biopsy (FNAB) represents the most important initial test for diagnosing malignancy. The result of the FNAB cytology can be classified as benign (70% of cases), malignant (5% to 10%), indeterminate or suspicious (10% to 20%), or nondiagnostic (10% to 15%).[3-5] Although nondiagnostic FNABs can be repeated, the indeterminate or suspicious group presents a dilemma for the clinician. In a recent report from our center on 80 patients who underwent thyroid resection for an indeterminate FNAB diagnosis of follicular neoplasm (FN), only 20% were confirmed as malignant.[6] Thus, many patients undergo thyroid surgery for nodular disease that is eventually diagnosed as benign disease.

Given the diagnostic limitations of FNAB when applied to thyroid tumors, multiple investigators have carried out expression profiling studies with hopes of identifying new diagnostic tools. Such analyses attempt to identify differentially expressed genes with an important role in disease development or progression using large-scale transcript-level expression profiling technologies such as cDNA microarrays,[7] oligonucleotide arrays[8] and Serial Analysis of Gene Expression (SAGE).[9] Typically, dozens or hundreds of genes are identified, many of which are expected to be false positives, and only a small fraction useful as diagnostic/prognostic markers or therapeutic targets.

A logical approach to distinguishing important genes from spurious genes, given a large number of candidate gene lists, is to search for the intersection of genes identified in multiple independent studies.[10] It is expected that biologically relevant genes will be over-represented and system-specific spurious genes under-represented. As large numbers of cancer profiling studies have become available, the identification of such intersections has become increasingly popular[10-12] but none have investigated thyroid cancer specifically. Such studies, although conceptually simple, face a number of technical challenges such as inconsistent gene identifiers, inaccessible data, and uncertain significance of results. Here, we attempt to overcome these challenges.

Our approach involves a vote-counting strategy based on the number of studies reporting a gene as differentially expressed and further ranking based on total sample size and average fold-change. Similar strategies have been used to show that gene pairs consistently coexpressed in multiple platforms[13] or data sets[14] are more likely to share a common biologic process. Our objective was to use validation from multiple expression profiling data sets to identify high-confidence, differentially expressed genes as potential biomarkers for thyroid cancer. We present a novel meta-review method for ranking genes on the basis of published evidence, successfully validate our method against a more traditional meta-analysis approach, and provide a large number of highly significant multistudy genes. Such markers should prove a useful resource for further study by high-throughput molecular analytic techniques.

Table 1. List of Abbreviations for Thyroid Samples

Abbreviation	Sample Description
ACL	Anaplastic thyroid cancer cell line
AFTN	Autonomously functioning thyroid nodules
ATC	Anaplastic thyroid cancer
CTN	Cold thyroid nodule
DTC	Differentiated thyroid cancer
FA	Follicular adenoma
FCL	Follicular carcinoma cell line
FTC	Follicular thyroid carcinoma
FVPTC	Follicular variant papillary thyroid carcinoma
GT	Goiter
HCC	Hurthle cell carcinoma
HN	Hyperplastic nodule
M	Metastatic
MACL	Anaplastic thyroid cancer cell line with metastatic capacity
MTC	Medullary thyroid carcinoma
Norm	Normal
PCL	Papillary carcinoma cell line
PTC	Papillary thyroid carcinoma
TCVPTC	Tall-cell variant papillary thyroid carcinoma
UCL	Undifferentiated carcinoma cell line

MATERIALS AND METHODS

Data Collection and Curation

Published lists of differentially expressed genes were processed to obtain the following information (wherever possible): unique identifier (probe/tag/accession); gene name/description; gene symbol; comparison conditions; sample numbers for each condition; fold-change; direction of change; and PubMed ID. All abbreviations used for sample descriptions are defined in Table 1.

Gene Mapping

The National Center for Biotechnology Information's Entrez gene identifier was chosen as the common target identifier for the overlap analysis. SAGE tags were mapped to transcripts by the first position (3′-most NlaIII anchoring enzyme recognition site), sense-strand tag predicted from Refseq[15] or MGC[16] sequences and then mapped to Entrez using the DiscoverySpace software package (Varhol et al, unpublished data; http://www.bcgsc.ca/discoveryspace/). Affymetrix probes were mapped using Affymetrix annotation files (Santa Clara, CA). Clone accession ids were mapped using the DAVID Resource (http://david.abcc.ncifcrf.gov/).[17] If no tag, probe, or accession ID was available, the entry was mapped using gene symbol or gene synonyms.

Ranking

Each published study consists of one or more comparisons between a pair of conditions (eg, papillary thyroid carcinoma [PTC] *v* normal) resulting in a list of differentially expressed genes. A method of ranking potential molecular markers was devised for each comparison group. A comparison group refers to a list of comparisons that address a common question of interest. For example, to identify markers that consistently distinguish cancer from noncancer (normal or benign) we would analyze all the comparisons that contrast cancer samples (eg, PTC, follicular thyroid carcinoma [FTC], anaplastic thyroid cancer [ATC], etc) against noncancer samples (eg, normal, goiter [GT], follicular adenoma [FA], etc).

Genes were ranked according to several criteria in the following order of importance: (1) number of comparisons in agreement (ie, listing the same gene as differentially expressed and with a consistent direction of change); (2) total number of samples for comparisons in agreement; and (3) average fold-change reported for comparisons in agreement. Total sample size was considered more important than average fold-change because many studies do not report a fold-change. Therefore, average fold-change was based solely on the subset of studies for which a fold-change value was available.

Assessment of Significance

Significance of the observed level of overlap between studies for each comparison subset was assessed by Monte Carlo simulation using custom Perl scripts. Where possible, the actual gene lists produced by mapping each expression technology to Entrez gene ID were utilized. For studies with custom arrays,[18-21] the appropriate number of genes was chosen from the combined gene list of all other platforms. For SAGE, three thyroid libraries (normal, benign, and carcinoma) from the Cancer Genome Anatomy Project[22] were used to create a realistic total tag set and then mapped to Entrez as noted herein. Once total gene lists were created for each platform type, we randomly created gene subsets of the same size observed in our review of the literature. For example, in the cancer-versus-noncancer analysis, one comparison (PTC *v* normal) identified 24 up- and 27 downregulated genes with the Affymetrix HG-U95A platform.[23] In our simulation, we would randomly select and label 24 "up" and 27 "down" genes from the Affymetrix HG-U95A total gene list. A similar random selection was performed for all other comparisons in the cancer-versus-noncancer subset using the appropriate total gene lists. Finally, the amount of overlap between comparisons was tallied as in the real analysis. This entire process was repeated 10,000 times to produce a distribution of overlap results from the random simulations. A P value was then estimated by comparing the actual overlap result to the distribution. A result was considered significant at $P < .05$.

Meta-Analysis of Affymetrix Data

The method presented in the preceding section makes use of reported lists of differentially expressed genes from published literature. An obvious disadvantage of this approach is that each publication may make use of different methods to ascertain differential expression (eg, scaling, filtering, normalization, significance thresholds, P value estimation, multiple testing corrections, etc). Collecting and reanalyzing 21 sets of raw data from 10 different platforms in a consistent manner would be an immense task and most likely impossible, because many raw data sets are unavailable. However,

to assess our method, we did reanalyze a subset of data from raw image files using a standard methodology. Five Affymetrix comparisons (three PTC *v* normal; one FTC *v* normal; and one FTC *v* FA) were reprocessed using the DChip software, analyzed for overlapping genes as above, and the results compared to the cancer-versus-noncancer comparison analysis for concordance using the LOLA tool.[11]

Additional information on methods appears in the Appendix (online only).

RESULTS

A total of 34 comparisons were available from 21 studies, utilizing 10 different expression platforms (Table 2). Of the 1,785 genes reported as differentially expressed in these studies (827 up- and 958 down-regulated), 1,562 could be mapped to an Entrez gene identifier (723

Table 2. Thyroid Cancer Profiling Studies Included in Analysis

Study	Platform	No. of Genes/ Features	Comparison				Features		Mapped Genes	
			Condition 1	No. of Samples	Condition 2	No. of Samples	Up-regulated	Down-regulated	Up-regulated	Down-regulated
Aldred et al, 2004[51]	Affymetrix HG-U95A	12,558	FTC	9	PTC	6	0	142	0	126
					Norm	13				
			PTC	6	FTC	9	68	0	59	0
					Norm	13				
Arnaldi et al, 2005[18]	Custom cDNA array	1,807	FCL	1	Norm	1	9	20	9	17
			FCL	1	Norm	1	3	6	3	3
			PCL	1						
			UCL	1						
			PCL	1	Norm	1	1	8	1	8
			UCL	1	Norm	1	1	7	1	6
Barden et al, 2003[49]	Affymetrix HG-U95A	12,558	FTC	9	FA	10	59	45	53	42
Cerutti et al, 2004[52]	SAGE	NA	FA	1	FTC	1	5	0	4	0
					Norm	1				
			FTC	1	FA	1	12	0	9	0
					Norm	1				
Chen et al, 2001[50]	Atlas human cDNA array (Clontech)	588	M	1	FTC	1	18	40	17	40
Chevillard et al, 2004[19]	Custom cDNA array	5,760	FTC	3	FA	4	12	31	12	30
			FVPTC	3	PTC	2	123	16	123	16
Eszlinger et al, 2001[53]	Atlas human cDNA array (Clontech)	588	AFTN	3	Norm	6	0	16	0	12
			CTN	3						
Finley et al, 2004[29]	Affymetrix HG-U95A	12,558	PTC	7	FA	14	48	85	48	82
			FVPTC	7	HN	7				
Finley et al, 2004[46]	Affymetrix HG-U95A	12,558	FTC	9	FA	16	50	55	49	52
			PTC	11	HN	10				
			FVPTC	13						
Giordano et al, 2005[48]	Affymetrix HG-U133A	22,283	PTC	51	Norm	4	90	151	69	122
Hawthorne et al, 2004[31]	Affymetrix HG-U95A	12,558	GT	6	Norm	6	1	7	0	6
			PTC	8	GT	6	10	28	8	18
			PTC	8	Norm	8	4	4	3	3
Huang et al, 2001[23]	Affymetrix HG-U95A	12,558	PTC	8	Norm	8	24	27	24	27
Jarzab et al, 2005[43]	Affymetrix HG-U133A	22,283	PTC	16	Norm	16	75	27	71	26
Mazzanti et al, 2004[55]	Hs-UniGem2 human cDNA array	9,984	PTC	17	FA	16	5	41	4	35
			FVPTC	15	HN	15				
Onda et al, 2004[20]	Amersham custom cDNA array	27,648	ACL	11	Norm	10	31	56	27	54
			ATC	10						
Pauws et al, 2004[57]	SAGE	NA	FVPTC	1	Norm	1	33	9	14	4
Takano et al, 2000[56]	SAGE	NA	FTC	1	ATC	1	3	10	1	7
			FTC	1	FA	1	4	1	2	1
			Norm	1	FA	1	6	0	2	0
			PTC	1	ATC	1	2	11	0	8
			PTC	1	FA	1	7	0	2	0
			PTC	1	FTC	1	2	1	1	1
Wasenius et al, 2003[35]	Atlas human cancer cDNA array (cancer 1.2 array)	1,176	PTC	18	Norm	3	12	9	12	8
Weber et al, 2005[47]	Affymetrix HG-U133A	22,283	FA	12	FTC	12	12	84	12	65
Yano et al, 2004[21]	Amersham custom cDNA array	3,968	PTC	7	Norm	7	54	0	41	0
Zou et al, 2004[54]	Atlas human cancer cDNA array (cancer 1.2 array)	1,176	MACL	1	ACL	1	43	21	42	20
21 studies	10 platforms		34 comparisons (473 samples)				827	958	723	839

NOTE. Table 1 contains definitions of thyroid sample abbreviations.
Abbreviations: Clontech, Clontech Laboratories Inc (Mountain View, CA); Amersham, Amersham Biosciences (Piscataway, NJ); Affymetrix, Affymetrix Inc (Santa Clara, CA); SAGE, Serial Analysis of Gene Expression; NA, not applicable.

up- and 839 downregulated). In all overlap analysis groups considered except for one, we identified genes that were reported in multiple studies with a level of overlap found to be significant by Monte Carlo simulation ($P < .05$; Table 3). The cancer-versus-noncancer group is provided as an example. In this case, a total of 755 genes were reported from 21 comparisons, and of these, 107 genes were reported more than once with consistent fold-change direction (Fig 1). In some cases (*MET*, *TFF3*, and *SERPINA1*), genes were independently reported as many as six times with a consistent fold-change direction. Only 18 genes were found to be reported in multiple studies with inconsistent fold-change. This in itself is an encouraging result. Given that approximately equal numbers of genes were reported as up- versus downregulated (723 up, 839 down) we might expect that multistudy genes with inconsistent fold-change direction would be as common as (or more common than) genes with consistent direction (under random expectation). Instead, we see that in most cases (85.6%), studies that report the same gene agree on the direction, even for large numbers of studies.

The total amount of overlap observed was assessed by Monte Carlo simulation and found to be highly significant ($P < .0001$; 10,000 permutations). In the simulation, an average of 18.2 (95% CI, 18.12 to 18.28) genes were observed with an overlap of two (same gene identified in two comparisons) compared with 68 in the actual data. For overlap of three, only 0.3 (95% CI, 0.29 to 0.31) genes were observed on average compared with 27 for real data. In 10,000 permutations, the simulated data never produced an overlap greater than three, whereas real data identified 12 genes with overlap of four, five, or six. The probability of observing one or more genes with an overlap of two or more was $P = .99$. For overlap of three or more $P = .037$, and for four or more $P < .0001$. The total number of genes with overlap of two was still highly significant, but we expect at least some false positives to occur by chance. Therefore, we have provided only those genes (top 39) with overlap of three or more and consider those with four or more to be the most reliable (Table 4).

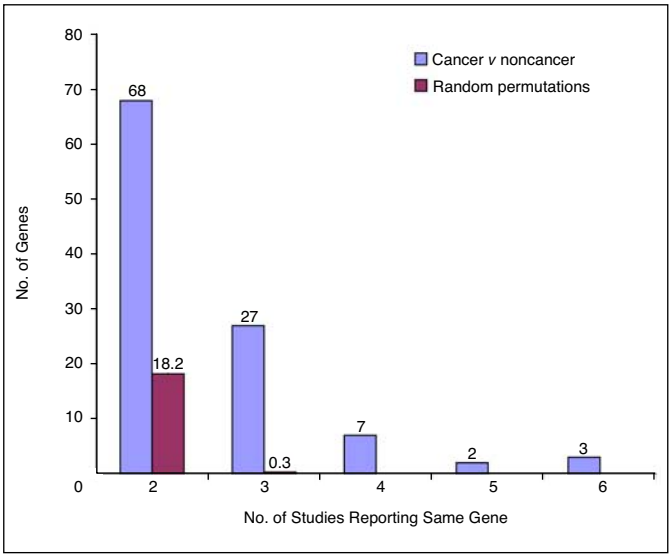

Fig 1. Overlap analysis results for cancer-versus-noncancer group compared with random simulation. Values shown for random permutations are mean values for all permutations in the Monte Carlo simulation. Error bars were not included because SE or 95% CIs were too small to visualize.

If the cancer-versus-noncancer group is broken into two categories, cancer versus normal and cancer versus benign, we find that most of the top genes were found in both types of comparisons. A small number of genes were found in only one of the two categories.

A comparison of genes with multistudy evidence based on published lists versus the smaller subset reanalyzed from raw Affymetrix

Table 3. Comparison Groups Analyzed for Overlap

Overlap Analysis Group	Condition Set 1	Condition Set 2	No. of Comparisons	Genes		*P*
				Total No.	No. With Multistudy Confirmation	
Cancer *v* noncancer	ACL, ATC, FCL, FTC, FVPTC, HCC, M, MACL, PCL, PTC, TCVPTC, UCL	AFTN, CTN, FA, GT, HN, Norm	21	755	107	< .0001
Cancer *v* normal	ACL, ATC, FCL, FTC, FVPTC, HCC, M, MACL, PCL, PTC, TCVPTC, UCL	Norm	12	478	53	< .0001
Cancer *v* benign	ACL, ATC, FCL, FTC, FVPTC, HCC, M, MACL, PCL, PTC, TCVPTC, UCL	AFTN, CTN, FA, GT, HN	8	332	38	< .0001
Normal *v* benign	Norm	AFTN, CTN, FA, GT, HN	3	19	1	0.0113
Papillary cancer *v* noncancer	FVPTC, PCL, PTC, TCVPTC	AFTN, CTN, FA, GT, HN, Norm	12	503	82	< .0001
Papillary cancer *v* normal	FVPTC, PCL, PTC, TCVPTC	Norm	8	369	49	< .0001
Papillary cancer *v* benign	FVPTC, PCL, PTC, TCVPTC	AFTN, CTN, FA, GT, HN	4	183	13	< .0001
Papillary cancer *v* other	FVPTC, PCL, PTC, TCVPTC	Any other	15	528	107	< .0001
FVPTC *v* other	FVPTC	Any other	2	157	0	NA
FTC *v* FA	FTC	FA	6	222	3	.0455
Follicular cancer *v* other	FTC, FCL	Any other	10	403	15	.0003
Aggressive cancer *v* other	ACL, ATC, M, MACL	Any other	4	145	4	.0402
Anaplastic cancer *v* other	ACL, ATC, MACL	Any other	3	91	6	< .0001
Cancer *v* noncancer (reanalyzed Affymetrix subset)	PTC, FTC	Norm, FA	5	1,317	179	< .0001

NOTE. Table 1 contains definitions of thyroid sample abbreviations.
Abbreviation: Affymetrix, Affymetrix Inc (Santa Clara, CA).

Table 4. Cancer Versus Noncancer (normal/benign) Overlap Analysis Results

Gene	Description	No. of Comparisons			No. of Samples		FC		References
		Upregulated	Downregulated	With FC	Total	With FC	Mean	Range	
MET‡	Met proto-oncogene (hepatocyte growth factor receptor)	6	0	4	202	162	4.54	2.60 to 6.60	23,35,43,46,48,49
TFF3‡	Trefoil factor 3 (intestinal)	0	6	4	196	146	−22.04	−63.55 to −3.80	20,23,31,46,48,49
SERPINA1‡	Serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	6	0	6	192	192	15.84	5.00 to 27.64	23,31,43,46,48
EPS8‡	Epidermal growth factor receptor pathway substrate 8	5	0	5	186	186	3.14	2.10 to 3.80	23,43,46-48
TIMP1‡	Tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	5	0	5	142	142	5.37	3.20 to 10.31	31,35,43,46
TGFA‡	Transforming growth factor, alpha	4	0	3	165	146	6.18	3.20 to 7.91	43,46,48,49
QPCT‡	Glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	4	0	4	153	153	7.31	3.40 to 11.67	19,43,46,48
PROS1‡	Protein S (alpha)	4	0	3	149	130	5.76	3.40 to 7.39	23,46,48,49
CRABP1‡	Cellular retinoic acid binding protein 1	0	4	4	146	146	−11.54	−24.45 to −2.20	23,31,46,48
FN1‡	Fibronectin 1	4	0	4	128	128	7.67	5.20 to 10.30	23,35,43,46
FCGBP‡	Fc fragment of IgG binding protein	0	4	3	108	89	−3.20	−3.30 to −3.10	31,46,49
TPO‡	Thyroid peroxidase	0	4	3	91	89	−6.25	−8.60 to −2.70	23,31,46,57
LRP4‡	Low-density lipoprotein receptor-related protein 4	3	0	3	146	146	14.47	6.40 to 19.43	43,46,48
PSD3‡	Pleckstrin and Sec7 domain containing 3	3	0	3	146	146	3.99	2.70 to 5.50	43,46,48
C11orf8‡	Chromosome 11 open reading frame 8	0	3	3	134	134	−7.04	−12.49 to −2.25	23,48,55
FABP4‡	Fatty acid binding protein 4, adipocyte	0	3	3	130	130	−8.55	−15.36 to −4.90	23,46,48
RGS16‡	Regulator of G-protein signaling 16	0	3	3	130	130	−4.01	−6.75 to −2.00	23,46,48
SDC4‡	Syndecan 4 (amphiglycan, ryudocan)	3	0	3	130	130	3.32	2.30 to 4.17	23,46,48
COL9A3‡	Collagen, type IX, alpha 3	0	3	3	128	128	−13.97	−27.39 to −4.50	31,46,48
HBB[*]	Hemoglobin, beta	0	3	2	118	87	−7.58	−11.39 to −3.77	20,43,48
ETV5‡	ets variant gene 5 (ets-related molecule)	3	0	3	111	111	3.60	2.98 to 4.38	43,47,48
CD44‡	CD44 antigen (homing function and Indian blood group system)	3	0	3	111	111	3.12	2.24 to 4.51	43,47,48
FCGRT‡	Fc fragment of IgG, receptor, transporter, alpha	0	3	1	109	59	−2.8	−2.80 to −2.80	20,46,49
CITED1‡	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 1	3	0	3	107	107	18.73	7.90 to 26.90	23,43,46
KRT19‡	Keratin 19	3	0	3	107	107	6.55	4.00 to 9.35	23,43,46
GPR51‡	G protein-coupled receptor 51	3	0	3	107	107	5.67	3.30 to 8.26	31,43,46
LGALS3‡	Lectin, galactoside-binding, soluble, 3 (galectin 3)	3	0	3	107	107	3.7	3.50 to 3.80	23,43,46
DPP4[*]	Dipeptidylpeptidase 4 (CD26, adenosine deaminase complexing protein 2)	3	0	3	103	103	46.19	8.20 to 115.76	23,43,48
TUSC3[*]	Tumor suppressor candidate 3	3	0	3	103	103	5.84	2.43 to 7.70	23,43,48
P4HA2[*]	Procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide II	3	0	3	103	103	3.75	2.93 to 4.50	23,43,48
CCND1[*]	cyclin D1 (PRAD1: parathyroid adenomatosis 1)	3	0	2	101	87	2.93	2.49 to 3.37	21,43,48
DIO1‡	Deiodinase, iodothyronine, type I	0	3	2	94	75	−3.75	−5.20 to −2.30	23,46,49
ITPR1‡	Inositol 1,4,5-triphosphate receptor, type 1	0	3	2	94	75	−2.6	−2.70 to −2.50	23,46,49

(continued on following page)

Table 4. Cancer Versus Noncancer (normal/benign) Overlap Analysis Results (continued)

Gene	Description	No. of Comparisons			No. of Samples		FC		References
		Upregulated	Downregulated	With FC	Total	With FC	Mean	Range	
MT1F†	Metallothionein 1F (functional)	0	3	2	92	73	−2.85	−2.91 to −2.80	31,46,49
PHLDA2‡	Pleckstrin homology-like domain, family A, member 2	3	0	3	89	89	8.02	2.50 to 15.96	23,31,46
MT1G‡	Metallothionein 1G	0	3	3	89	89	−5.55	−8.60 to −2.30	23,31,46
ID4†	Inhibitor of DNA binding 4, dominant negative helix-loop-helix protein	0	3	2	89	70	−3.64	−5.00 to −2.29	19,49,55
DUSP6‡	Dual-specificity phosphatase 6	3	0	3	72	72	4.5	3.68 to 5.22	23,43,47
HBA2[*]	Hemoglobin, alpha 2	0	3	3	69	69	−3.5	−4.72 to −2.38	23,35,43

Abbreviation: FC, fold change.
[*]Cancer *v* normal.
†Cancer *v* benign.
‡Both.

microarray data showed a highly significant level of agreement ($P < .0001$). The 107 cancer-versus-noncancer multistudy genes showed a concordance of 0.177 (95% CI, 0.129 to 0.225) with the 179 multistudy genes identified from the reanalyzed Affymetrix subset (Fig 2). In total, there were 43 genes identified by both methods. Given that the two lists of genes were produced by very different subsets of data, in addition to the potential differences in processing, this was an encouraging result. However, it does appear that reprocessing the microarray data in a consistent manner would certainly alter the results and would likely increase the total number of multistudy genes.

Additional information on results appears in the Appendix (online only).

DISCUSSION

A common criticism of expression profiling studies is a lack of agreement between studies. However, by applying our meta-review method to a large number of published studies, we observe that many genes are consistently reported at a highly significant rate. These genes may represent real biologic effects that, through repeated efforts, have overcome the issues of noise and error typically associated with such experiments. A comparison of our meta-review method (using published gene lists) to a meta-analysis of a smaller subset of studies (for which raw data were available) showed a strong level of concordance. Thus, we believe our approach represents a useful alternative for identifying consistent gene expression markers when raw data are unavailable (as is generally the case). However, a limitation of our method resulting from unavailability of raw data is that we are unable to assign a measure of confidence at the gene level. We can identify consistently reported genes and rank them according to simple criteria such as total sample size and average fold-change, but we can not calculate a true combined fold-change or *P* value. In order for more powerful meta-analysis methods to be applied researchers must provide access to their raw data. Also, we remind the reader that although we have focused on the cancer-versus-noncancer comparisons, a large number of other comparison groups were analyzed (Table 3).

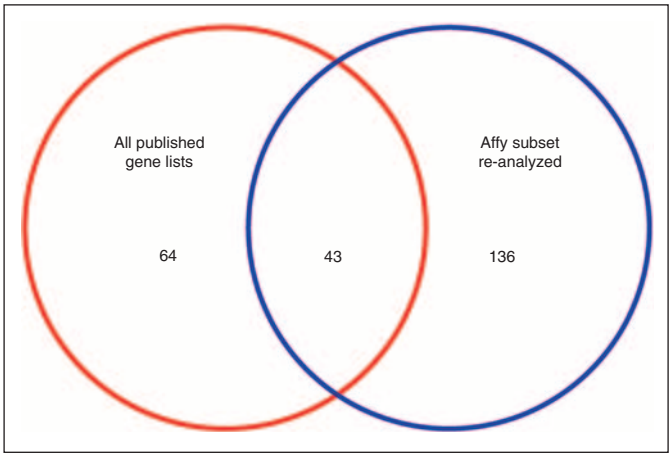

Fig 2. A comparison of cancer-versus-noncancer genes identified with multistudy evidence based on all published lists (our meta-review method) versus genes identified by a smaller subset of studies reanalyzed from raw microarray data. Affy, Affymetrix, Santa Clara, CA.

As a means of further assessing our results, we review the top 12 cancer-versus-noncancer candidates to identify which markers have been previously confirmed as differentially expressed or having diagnostic/prognostic utility in thyroid cancer (Table 5). In total, 10 of 12 markers have been confirmed at the RNA level and six of these have gone on to be validated at the protein level. For discussion purposes we have broken the genes into two categories, well-characterized and novel or uncharacterized. We also compare our results to a previous review of promising thyroid biomarkers.

We defined well-characterized genes as those that have been validated in more than one follow-up study and at both the RNA and protein level, such as *MET*, *TFF3*, *SERPINA1*, *TIMP1*, *FN1*, and *TPO*. Several studies have implicated MET protein expression in thyroid cancer as both a diagnostic tool[24-28] and prognostic tool.[24,26-28] Increased MET expression has been associated with higher risk for metastasis[26] and recurrence in PTC[26,27] and negative prognosis in FTC.[28] However, in another study, decreased MET was shown to be an effective predictor of distant metastases among PTC cases.[24] Although no reports have evaluated TFF3 at the protein level, numerous studies have suggested TFF3 as a useful biomarker at the RNA level.[23,29-33] A two-gene panel of *SFTPB* and *TFF3* was shown to correctly diagnose

Table 5. Summary of Experimental Validation for Top 12 Markers

Gene	RNA	Protein
MET	RT-PCR[29,35,49,58]	IHC, [24,26,27,59] IS,[25] WB[58,60]
TFF3	RT-PCR[23,29-33]	
SERPINA1	RT-PCR[31]	IHC, [34] WB[34]
EPS8		
TIMP1	RT-PCR,[31,35] NB[61]	IHC[35,36,61]
TGFA	NB[42]	ELISA[42]
QPCT	RT-PCR[43]	
PROS1		
CRABP1	RT-PCR[31]	
FN1	RT-PCR[23,35,37]	IHC[35,39]
FCGBP	RT-PCR[45]	
TPO	RT-PCR,[21,40,57] NB[21,62]	IHC[41,48]

Abbreviations: RT-PCR, reverse transcriptase polymerase chain reaction; IHC, immunohistochemistry; IS, immunostaining; NB, northern blot; WB, western blot; ELISA, enzyme-linked immunosorbent assay.

PTC with a sensitivity of 88.9%, specificity of 96.7%, and accuracy of 94.9%[30] and *TFF3*/*LGALS3* mRNA ratio was shown to distinguish FA from FTC with sensitivity and specificity of 80.0% and 91.5% respectively.[33] An antibody study of *SERPINA1* reported immunoreactivity in nine of 10 PTCs with no staining in the adjacent normal thyroid tissues.[34] *TIMP1* upregulation was confirmed by immunohistochemistry (IHC) with positive immunostaining in 68% of PTC cases and none of the normal cases.[35] Another IHC study of *TIMP1* for 86 PTC specimens showed increased immunoreactivity in the tumor regions versus nontumor regions in 92% cases and significant correlations with unfavorable prognostic variables.[36] *FN1* has been proposed as a useful reverse transcriptase (RT-) polymerase chain reaction (PCR) marker of differentiated thyroid cancer (DTC)[37] and an important modulator of thyroid cell adhesiveness and neoplastic cell growth.[38] An IHC study of 85 FTCs and 21 FAs reported that coexpression of *FN1* and *GAL3* or *FN1* and *HBME1* was restricted to cancer, although their concurrent absence was highly specific for benign lesions (96%).[39] A large number of studies have investigated *TPO* as a marker for thyroid carcinoma. Lazar et al[40] found that higher thyroid cancer stage was associated with lower *TPO* mRNA expression. Segev et al[41] reviewed five IHC studies involving nearly 400 follicular lesions and found that 93% of FAs and 97% of FTCs were accurately diagnosed by *TPO* antibody staining. Studies using FNAB samples, however, have proved less promising with false-positive rates as high as 32%.[41] For the most part, the six genes reviewed above appear promising as thyroid cancer candidates and suggest our meta-analysis method is producing reasonable results.

For four genes (*TGFA*, *QPCT*, *CRABP1*, and *FCGBP*) we could find only a single follow-up study or validation experiment confirming their potential importance in thyroid cancer. Bergstrom et al[42] suggest that increased expression of *TGFA* may be responsible for aberrant activation of epidermal growth factor receptor and ultimately an overexpression and activation of *MET*. Jarzab et al[43] built a classifier capable of discriminating between PTC and nonmalignant samples in 90% of cases. This classifier included *QPCT* (along with 18 other genes). *QPCT* was considered a novel gene and was validated by quantitative PCR in that study, but has been studied little further since. *CRABP1* downregulation was confirmed by RT-PCR (in one of the original microarray studies),[31] and another study reported that hypermethylation of promoter CpG islands for *CRABP1* in PTC may explain the reduced expression.[44] Differential expression of *FCGBP* was confirmed in a separate study by restriction-mediated differential display and real-time RT-PCR.[45]

For two genes (*EPS8* and *PROS1*) we could find no confirmation beyond the initial microarray experiment. In our meta-analysis, five studies identified *EPS8*[23,43,46-48] and four identified *PROS1*[23,46,48,49] as upregulated in comparisons of cancer with noncancer. And yet, to our knowledge, no follow-up study has confirmed either of these genes (even at the RNA level). It is unclear whether genes such as *EPS8* and *PROS1* have not been further validated because they are false positives or simply because they have not yet been chosen for further study. These genes and the other less characterized candidates may represent novel diagnostic markers for thyroid cancer and warrant further investigation.

Comparison to a previous meta-review by Segev et al[41] of mainly single-gene, protein-level thyroid cancer studies found that four of their 12 markers identified as promising preoperative diagnostic markers were identified as high-ranking candidates (top 30) in our meta-analysis (*TPO*, *CD44*, *KRT19*, and *LGALS3*). Two of their candidates were either not represented (*HBME-1*) or can not be reliably assayed by the microarray platforms (*RET/PTC* rearrangements). However, six other promising markers (*CDKN1B*, *TERT*, *CP/LTF*, *DLGAP4*, *HMGA1*, and *PAX8*) do have representation on at least some of the expression platforms, and yet were not identified as differentially expressed in even a single study in our meta-analysis. These genes may have displayed some differential expression but not reached the required thresholds for inclusion in the published lists. Or, they may represent cases in which changes in RNA levels do not correlate well with changes in protein levels. Segev et al[41] concluded that large-scale thyroid tumor expression profiling of multiple markers on tumors from large and diverse patient cohorts are still required to identify a panel of markers with sufficient sensitivity and specificity to accurately diagnose indeterminate thyroid lesions. We agree and believe that our meta-review of thyroid cancer gene expression profiling studies provides a high-quality list of candidates from which to identify such a panel.

REFERENCES

1. Gharib H, Goellner JR: Fine-needle aspiration biopsy of the thyroid: An appraisal. Ann Intern Med 118:282-289, 1993

2. Greenlee RT, Hill-Harmon MB, Murray T, et al: Cancer statistics, 2001. CA Cancer J Clin 51:15-36, 2001

3. Goellner JR, Gharib H, Grant CS, et al: Fine needle aspiration cytology of the thyroid, 1980 to 1986. Acta Cytol 31:587-590, 1987

4. Caraway NP, Sneige N, Samaan NA: Diagnostic pitfalls in thyroid fine-needle aspiration: A review of 394 cases. Diagn Cytopathol 9:345-350, 1993

5. Ravetto C, Colombo L, Dottorini ME: Usefulness of fine-needle aspiration in the diagnosis of thyroid carcinoma: A retrospective study in 37,895 patients. Cancer 90:357-363, 2000

6. Wiseman SM, Baliski C, Irvine R, et al: Hemithyroidectomy: The optimal initial surgical approach for individuals undergoing surgery for a cytological diagnosis of follicular neoplasm. Ann Surg Oncol 13:425-432, 2006

7. Schena M, Shalon D, Davis RW, et al: Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270: 467-470, 1995

8. Lockhart DJ, Dong H, Byrne MC, et al: Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol 14:1675-1680, 1996

9. Velculescu VE, Zhang L, Vogelstein B, et al: Serial analysis of gene expression. Science 270:484-487, 1995

10. Rhodes DR, Yu J, Shanker K, et al: Large-scale meta-analysis of cancer microarray data identifies

common transcriptional profiles of neoplastic transformation and progression. Proc Natl Acad Sci U S A 101:9309-9314, 2004

11. Cahan P, Ahmad AM, Burke H, et al: List of lists-annotated (LOLA): A database for annotation and comparison of published microarray gene lists. Gene 360:78-82, 2005

12. Shih W, Chetty R, Tsao MS: Expression profiling by microarrays in colorectal cancer. Oncol Rep 13:517-524, 2005

13. Griffith OL, Pleasance ED, Fulton DL, et al: Assessment and integration of publicly available SAGE, cDNA microarray, and oligonucleotide microarray expression data for global coexpression analyses. Genomics 86:476-488, 2005

14. Lee HK, Hsu AK, Sajdak J, et al: Coexpression analysis of human genes across many microarray data sets. Genome Res 14:1085-1094, 2004

15. Pruitt KD, Katz KS, Sicotte H, et al: Introducing RefSeq and LocusLink: Curated human genome resources at the NCBI. Trends Genet 16:44-47, 2000

16. Mammalian Gene Collection Program Team, Strausberg RL, Feingold EA, et al: Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U S A 99:16899-16903, 2002

17. Dennis G Jr, Sherman BT, Hosack DA, et al: DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4:P3, 2003

18. Arnaldi LA, Borra RC, Maciel RM, et al: Gene expression profiles reveal that DCN, DIO1, and DIO2 are underexpressed in benign and malignant thyroid tumors. Thyroid 15:210-221, 2005

19. Chevillard S, Ugolin N, Vielh P, et al: Gene expression profiling of differentiated thyroid neoplasms: Diagnostic and clinical implications. Clin Cancer Res 10:6586-6597, 2004

20. Onda M, Emi M, Yoshida A, et al: Comprehensive gene expression profiling of anaplastic thyroid cancers with cDNA microarray of 25 344 genes. Endocr Relat Cancer 11:843-854, 2004

21. Yano Y, Uematsu N, Yashiro T, et al: Gene expression profiling identifies platelet-derived growth factor as a diagnostic molecular marker for papillary thyroid carcinoma. Clin Cancer Res 10: 2035-2043, 2004

22. Strausberg RL, Buetow KH, Greenhut SF, et al: The cancer genome anatomy project: Online resources to reveal the molecular signatures of cancer. Cancer Invest 20:1038-1050, 2002

23. Huang Y, Prasad M, Lemon WJ, et al: Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. Proc Natl Acad Sci U S A 98:15044-15049, 2001

24. Belfiore A, Gangemi P, Costantino A, et al: Negative/low expression of the Met/hepatocyte growth factor receptor identifies papillary thyroid carcinomas with high risk of distant metastases. J Clin Endocrinol Metab 82:2322-2328, 1997

25. Ippolito A, Vella V, La Rosa GL, et al: Immunostaining for Met/HGF receptor may be useful to identify malignancies in thyroid lesions classified suspicious at fine-needle aspiration biopsy. Thyroid 11:783-787, 2001

26. Ramirez R, Hsu D, Patel A, et al: Overexpression of hepatocyte growth factor/scatter factor (HGF/SF) and the HGF/SF receptor (cMET) are associated with a high risk of metastasis and recurrence for children and young adults with papillary thyroid carcinoma. Clin Endocrinol (Oxf) 53:635-644, 2000

27. Inaba M, Sato H, Abe Y, et al: Expression and significance of c-met protein in papillary thyroid carcinoma. Tokai J Exp Clin Med 27:43-49, 2002

28. Di Renzo MF, Olivero M, Ferro S, et al: Overexpression of the c-MET/HGF receptor gene in human thyroid carcinomas. Oncogene 7:2549-2553, 1992

29. Finley DJ, Arora N, Zhu B, et al: Molecular profiling distinguishes papillary carcinoma from benign thyroid nodules. J Clin Endocrinol Metab 89: 3214-3223, 2004

30. Hamada A, Mankovskaya S, Saenko V, et al: Diagnostic usefulness of PCR profiling of the differentially expressed marker genes in thyroid papillary carcinomas. Cancer Lett 224:289-301, 2005

31. Hawthorn L, Stein L, Varma R, et al: TIMP1 and SERPIN-A overexpression and TFF3 and CRABP1 underexpression as biomarkers for papillary thyroid carcinoma. Head Neck 26:1069-1083, 2004

32. Takano T, Miyauchi A, Yoshida H, et al: High-throughput differential screening of mRNAs by serial analysis of gene expression: Decreased expression of trefoil factor 3 mRNA in thyroid follicular carcinomas. Br J Cancer 90:1600-1605, 2004

33. Takano T, Miyauchi A, Yoshida H, et al: Decreased relative expression level of trefoil factor 3 mRNA to galectin-3 mRNA distinguishes thyroid follicular carcinoma from adenoma. Cancer Lett 219: 91-96, 2005

34. Poblete MT, Nualart F, del Pozo M, et al: Alpha 1-antitrypsin expression in human thyroid papillary carcinoma. Am J Surg Pathol 20:956-963, 1996

35. Wasenius VM, Hemmer S, Kettunen E, et al: Hepatocyte growth factor receptor, matrix metalloproteinase-11, tissue inhibitor of metalloproteinase-1, and fibronectin are up-regulated in papillary thyroid carcinoma: A cDNA and tissue microarray study. Clin Cancer Res 9:68-75, 2003

36. Maeta H, Ohgi S, Terada T: Protein expression of matrix metalloproteinases 2 and 9 and tissue inhibitors of metalloproteinase 1 and 2 in papillary thyroid carcinomas. Virchows Arch 438:121-128, 2001

37. Hesse E, Musholt PB, Potter E, et al: Oncofoetal fibronectin–a tumour-specific marker in detecting minimal residual disease in differentiated thyroid carcinoma. Br J Cancer 93:565-570, 2005

38. Liu W, Asa SL, Ezzat S: 1alpha,25-Dihydroxyvitamin D3 targets PTEN-dependent fibronectin expression to restore thyroid cancer cell adhesiveness. Mol Endocrinol 19:2349-2357, 2005

39. Prasad ML, Pellegata NS, Huang Y, et al: Galectin-3, fibronectin-1, CITED-1, HBME1 and cytokeratin-19 immunohistochemistry is useful for the differential diagnosis of thyroid tumors. Mod Pathol 18:48-57, 2005

40. Lazar V, Bidart JM, Caillou B, et al: Expression of the Na+/I- symporter gene in human thyroid tumors: A comparison study with other thyroid-specific genes. J Clin Endocrinol Metab 84:3228-3234, 1999

41. Segev DL, Clark DP, Zeiger MA, et al: Beyond the suspicious thyroid fine needle aspirate: A review. Acta Cytol 47:709-722, 2003

42. Bergstrom JD, Westermark B, Heldin NE: Epidermal growth factor receptor signaling activates met in human anaplastic thyroid carcinoma cells. Exp Cell Res 259:293-299, 2000

43. Jarzab B, Wiench M, Fujarewicz K, et al: Gene expression profile of papillary thyroid cancer: Sources of variability and diagnostic implications. Cancer Res 65:1587-1597, 2005

44. Huang Y, de la Chapelle A, Pellegata NS: Hypermethylation, but not LOH, is associated with the low expression of MT1G and CRABP1 in papillary thyroid carcinoma. Int J Cancer 104:735-744, 2003

45. O'Donovan N, Fischer A, Abdo EM, et al: Differential expression of IgG Fc binding protein (FcgammaBP) in human normal thyroid tissue, thyroid adenomas and thyroid carcinomas. J Endocrinol 174:517-524, 2002

46. Finley DJ, Zhu B, Barden CB, et al: Discrimination of benign and malignant thyroid nodules by molecular profiling. Ann Surg 240:425-437, 2004

47. Weber F, Shen L, Aldred MA, et al: Genetic classification of benign and malignant thyroid follicular neoplasia based on a 3-gene combination. J Clin Endocrinol Metab 90:2512-2521, 2005

48. Giordano TJ, Kuick R, Thomas DG, et al: Molecular classification of papillary thyroid carcinoma: Distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. Oncogene 24:6646-6656, 2005

49. Barden CB, Shister KW, Zhu B, et al: Classification of follicular thyroid tumors by molecular signature: Results of gene profiling. Clin Cancer Res 9:1792-1800, 2003

50. Chen KT, Lin JD, Chao TC, et al: Identifying differentially expressed genes associated with metastasis of follicular thyroid cancer by cDNA expression array. Thyroid 11:41-46, 2001

51. Aldred MA, Huang Y, Liyanarachchi S, et al: Papillary and follicular thyroid carcinomas show distinctly different microarray expression profiles and can be distinguished by a minimum of five genes. J Clin Oncol 22:3531-3539, 2004

52. Cerutti JM, Delcelo R, Amadei MJ, et al: A preoperative diagnostic test that distinguishes benign from malignant thyroid carcinoma based on gene expression. J Clin Invest 113:1234-1242, 2004

53. Eszlinger M, Krohn K, Paschke R: Complementary DNA expression array analysis suggests a lower expression of signal transduction proteins and receptors in cold and hot thyroid nodules. J Clin Endocrinol Metab 86:4834-4842, 2001

54. Zou M, Famulski KS, Parhar RS, et al: Microarray analysis of metastasis-associated gene expression profiling in a murine model of thyroid carcinoma pulmonary metastasis: Identification of S100A4 (Mts1) gene overexpression as a poor prognostic marker for thyroid carcinoma. J Clin Endocrinol Metab 89:6146-6154, 2004

55. Mazzanti C, Zeiger MA, Costouros NG, et al: Using gene expression profiling to differentiate benign versus malignant thyroid tumors. Cancer Res 64:2898-2903, 2004

56. Takano T, Hasegawa Y, Matsuzuka F, et al: Gene expression profiles in thyroid carcinomas. Br J Cancer 83:1495-1502, 2000

57. Pauws E, Veenboer GJ, Smit JW, et al: Genes differentially expressed in thyroid carcinoma identified by comparison of SAGE expression profiles. Faseb J 18:560-561, 2004

58. Scarpino S, Cancellario d'Alena F, Di Napoli A, et al: Increased expression of Met protein is associated with up-regulation of hypoxia inducible factor-1 (HIF-1) in tumour cells in papillary carcinoma of the thyroid. J Pathol 202:352-358, 2004

59. Nardone HC, Ziober AF, LiVolsi VA, et al: C-Met expression in tall cell variant papillary carcinoma of the thyroid. Cancer 98:1386-1393, 2003

60. Mineo R, Costantino A, Frasca F, et al: Activation of the hepatocyte growth factor (HGF)-Met

system in papillary thyroid cancer: Biological effects of HGF in thyroid cancer cells depend on Met expression levels. Endocrinology 145:4355-4365, 2004

61. Shi Y, Parhar RS, Zou M, et al: Tissue inhibitor of metalloproteinases-1 (TIMP-1) mRNA is elevated in advanced stages of thyroid carcinoma. Br J Cancer 79:1234-1239, 1999

62. Umeki K, Tanaka T, Yamamoto I, et al: Differential expression of dipeptidyl peptidase IV (CD26) and thyroid peroxidase in neoplastic thyroid tissues. Endocr J 43:53-60, 1996

Acknowledgment

We would like to acknowledge the generosity of the researchers who shared their data with us through personal communications including, Frank Weber, Carrie Ris-Stalpers, Charis Eng, Sandya Liyanarachchi, and Twyla Pohar.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Authors' Disclosures of Potential Conflicts of Interest

The authors indicated no potential conflicts of interest.

Author Contributions

Conception and design: Obi L. Griffith, Steven J.M. Jones, Sam M. Wiseman
Financial support: Steven J.M. Jones, Sam M. Wiseman
Collection and assembly of data: Obi L. Griffith, Sam M. Wiseman
Data analysis and interpretation: Obi L. Griffith, Adrienne Melck
Manuscript writing: Obi L. Griffith, Adrienne Melck, Sam M. Wiseman
Final approval of manuscript: Obi L. Griffith, Adrienne Melck, Steven J.M. Jones, Sam M. Wiseman