Ovarian Cancer Tumor Marker Behavior in Asymptomatic Healthy Women: Implications for Screening

Casey Crump,[1] Martin W. McIntosh, Nicole Urban, Garnet Anderson, and Beth Y. Karlan

Department of Biostatistics, University of Washington, Seattle, Washington 98195 [C. C., M. W. M., G. A.]; Cancer Prevention Research Program, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109 [M. W. M., N. U., G. A.]; and Department of Obstetrics and Gynecology at Cedars-Sinai Medical Center and the University of California, Los Angeles, California 90048 [B. Y. K.]

Abstract

Ovarian cancer screening protocols generally have been limited by inadequate recognition of the normal behavior of tumor markers in women at risk of ovarian cancer. We have characterized the behavior of five serum tumor markers in a large cohort of healthy women and examined the implications for screening. Serial measurements of CA125, HER-2/*neu*, urinary gonadotropin peptide, lipid-associated sialic acid, and Dianon marker 70/K were obtained during 6 years of follow-up of 1257 healthy women at high risk of ovarian cancer. We analyzed individual-specific tumor marker behavior and explored methods that can exploit this information to develop individual-specific screening rules. These five tumor markers behaved approximately independently. Substantial heterogeneity was observed among women in the behavior of each tumor marker, particularly CA125. Intraclass correlation (ICC), or the proportion of total variability that occurs between individuals, was approximately 0.6 for log-transformed CA125 and urinary gonadotropin peptide, and less than 0.4 for the other markers. This degree of tumor marker heterogeneity among healthy women, and the relative independence of these markers, has important implications for screening and diagnostic tests. Independence of markers results in the clinically useful fact that the combined false positive rate from screening with multiple markers may be estimated by the sum of individual false positive rates. Heterogeneity of tumor marker patterns in healthy women implies that a fixed screening cutoff level is suboptimal to a degree that depends strongly on ICC. Using information on longitudinal measurements and ICC, individual-specific screening rules may be developed with the potential to improve early detection of ovarian cancer.

Introduction

Ovarian cancer is the leading cause of death due to gynecological malignancy in the United States. This year an estimated 23,100 new cases of ovarian cancer will be diagnosed in the U.S., and 14,000 women will die from this disease (1). Most ovarian cancers occur in postmenopausal women, with over half occurring in women over the age of 60 years. Therapeutic advances have had a modest effect on the overall survival rate during the past three decades. Currently only 46% of women diagnosed with ovarian cancer will survive longer than 5 years. If disease is detected before dissemination beyond the ovaries, however, the 5-year survival rate is 93%. Unfortunately, only 24% of ovarian cancers are detected at this early stage (1). Development of an effective screening program for early detection, especially in high-risk women, is a top priority for reducing mortality due to ovarian cancer.

Research into noninvasive first-line screening techniques has focused on serum tumor markers and TVS.[2] TVS can detect ovarian cancer with a sensitivity approaching 100%, but it is insufficiently specific and too expensive for use as a first-line screen. A multimodal strategy using TVS only in women with a positive serum tumor marker screen appears to be the most promising alternative to lowering error rates and improving cost-effectiveness. Such an approach requires that all positive first-line tumor marker screenings be evaluated by TVS before a definitive surgical diagnosis is sought.

Ovarian cancers evolve from different cell types, and therefore may require different tumor-associated antigens for monitoring or diagnostic screening. Using monoclonal antibody assays, a large number of markers have been identified that are associated with epithelial ovarian carcinomas. CA125, the most extensively studied of these, is a high molecular weight glycoprotein expressed by most epithelial ovarian cancers. Elevations of CA125 in the serum may precede stage III ovarian cancer by 10 to 12 months (2), but the percentage of women with stage I disease who have an elevated CA125 level (>35 units/ml) ranges from only 23 to 50% (3, 4). CA125 has an established role in the monitoring of patients with known ovarian cancer, but in its present use has been insufficiently sensitive and specific to be very effective as a diagnostic screening tool for early-stage disease.

Other candidate markers with a potential role in ovarian cancer diagnostic screening include UGP, LASA, DM/70K, and HER-2/*neu*. In some populations UGP has been a sensitive marker for gynecological malignancies, with a specificity exceeding 90% for ovarian cancer (5). LASA is a marker for malignant disease on the basis of elevated sialoglycoconjugate levels associated with tumor growth (6), but it is relatively

Received 2/28/00; revised 7/7/00; accepted 7/21/00.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom requests for reprints should be addressed, at 5307 Ravenna Place NE, #3, Seattle, WA 98105. Phone: (206) 985-8615; E-mail: kccrump@u.washington.edu.

[2] The abbreviations used are: TVS, trans-vaginal sonography; UGP, urinary gonadotropin peptide; LASA, lipid-associated sialic acid; DM/70K, Dianon marker 70/K; ICC, intraclass correlation.

nonspecific for ovarian cancer (7). The DM/70K marker is immunologically distinct from CA125 and potentially a good adjunct; it was extracted from epithelial ovarian tumor tissue and can be found in patients with epithelial ovarian tumors of all histological types, including mucinous tumors (8). HER-2/*neu* encodes a transmembrane glycoprotein shed into the sera of some patients with ovarian cancer. Amplification of the HER-2/*neu* oncogene has been associated with poor survival in patients with ovarian cancer whose tumors overexpress this oncogene (9). HER-2/*neu* is amplified or overexpressed in 25 to 30% of ovarian carcinomas, though its usefulness for screening has not yet been demonstrated (10, 11).

Recent efforts to improve ovarian cancer screening have focused on modeling longitudinal tumor marker measurements, and on jointly monitoring CA125 with other markers. Skates *et al.* (12) proposed a screening algorithm based on a linear regression model of serial CA125 measurements which yielded a positive predictive value of 16%, substantially greater than that based on a single assay (<2%). This work helped demonstrate the substantial potential benefit of using longitudinal biomarker measurements in ovarian cancer screening. Algorithms that include multiple markers have yielded mixed results. A combination of CA125, M-CSF, and OVX-1 detected more than 95% of stage I ovarian cancers in a retrospective postmenopausal cohort, with specificity approaching 90% (13). Other studies have suggested that the joint elevation of CA125 with adjunct markers may provide a more specific test than CA125 alone, but at the expense of decreased sensitivity (14).

Cane *et al.* (15) described trends between first and second measurements of CA125, LASA, DM/70K, UGP, and HER-2/*neu* in 425 premenopausal and 165 postmenopausal healthy women enrolled in the Gilda Radner Ovarian Cancer Detection Program. These data suggested that plateaus of elevated values may be more frequent in healthy women than was previously believed. Although 80% of elevated first-screen LASA, CA125, and DM/70K markers returned to normal levels on a second screening, over half of elevated HER-2/*neu* and UGP values remained elevated. These findings suggested that current definitions of normal tumor marker values should be reevaluated.

If the normal behavior of tumor markers is heterogeneous among healthy women, the use of a fixed screening cutoff level will not optimally detect elevations from each individual's baseline. Most screening protocols to date have relied on a fixed cutoff for all individuals screened and thus have not taken advantage of information on heterogeneity of tumor marker levels among individuals [the algorithm by Skates *et al.* (12) is one exception]. To establish optimal screening rules for what is abnormal, a better recognition of the normal behavior of tumor markers in cancer-free women is clearly needed. The main objective of this analysis was to describe such behavior of five tumor markers in a large cohort of women at high risk of ovarian cancer.

Materials and Methods

Patient Population. The Gilda Radner Ovarian Cancer Detection Program began in 1991 at the Cedars-Sinai Medical Center in Los Angeles (7). In its first 6 years the program enrolled 1257 asymptomatic women volunteers from the Los Angeles area through physician- and self-referral. To be eligible, a woman had to be asymptomatic, available for follow-up for the next 5 years, and have a family history of either a first- or second-degree relative with ovarian cancer, or a first-degree relative with breast cancer, other gynecological cancer, or colon cancer. After obtaining informed consent, family histories were confirmed by pathology reports or death certificates. Screening with a battery of five serum tumor markers (CA125, HER-2/*neu*, UGP, LASA, and DM/70K), TVS, and color Doppler imaging was performed semiannually from July 1991 until July 1995, when the protocol was changed to annual screening. Tumor marker assays were performed in blinded fashion by Dianon Systems (Stratford, CT) using techniques described previously for CA125 (3), HER-2/*neu* (9), UGP (5), LASA (6), and DM/70K (8). In the first 6 years of follow-up, 10 ovarian cancer cases were detected (16). These cases were not the focus of and were not included in the present analysis.

Statistical Analysis. To assess for heterogeneity of tumor marker levels among women, ANOVA was used to estimate between-person variance (τ^2), within-person variance (σ^2), and intraclass correlation ($ICC = \tau^2/[\tau^2 + \sigma^2]$) separately for all five tumor markers. ICC is the proportion of total variability in tumor marker levels accounted for by variability among mean levels of different individuals. Heterogeneity of individual means was tested using the overall ANOVA F statistic, and heterogeneity of individual variances was tested using Bartlett's χ^2 statistic.

Empirical Bayes analysis was used to estimate individual-specific tumor marker means and standard deviations (17, 18). For the estimation of multiple means, empirical Bayes analysis gives better (least mean squared error) unbiased estimates than any other known method. Individual means are estimated as an empirical posterior distribution, using all available data from the cohort to form a prior distribution. The individual means are estimated intermediate between the arithmetic within-person mean and the overall cohort mean. The amount of "shrinkage" toward the overall mean is 1 minus a shrinkage factor which resembles the ICC and is a function of the number and variability of measurements observed for the individual (shrinkage factor: $B(n) = \tau^2/[\tau^2 + \sigma^2/\mathrm{n}]$). Empirical Bayes estimation is easily extended to multivariate models using two or more tumor markers. [See Efron and Morris (19) for a good nontechnical introduction to the empirical Bayes method.]

A hierarchical Bayes linear model (20) was used to characterize the bivariate behavior of CA125 and HER-2/*neu*. The hierarchical Bayes approach can be used to estimate individual-level means (μ_i) and standard deviations (σ_i) by combining multiple individuals' tumor marker distributions and to estimate the variability of μ_i in the population (τ) and the variability of σ_i in the population. This methodology can incorporate random variation at different levels (within-person [σ^2] and between-person [τ^2] variance) and the effect of specific factors (*e.g.*, menopausal status) that may cause heterogeneity in tumor marker patterns among women.

Results

Cohort Characteristics. From 1991 to 1997, serial measurements of CA125, LASA, DM/70K, UGP, and HER-2/*neu* were obtained from 1257 women who made a total of 5358 screening visits. On average, the number of visits/woman was 4.3 (ranging from 1 to 14), and the average interval between screenings was 7.8 months. Characteristics of the cohort are given in Table 1. Sixty-three percent of women were premenopausal (defined as <40 days since last menstrual period at all observations), 27.5% were postmenopausal (≥180 days since last menstrual period at all observations), and 8.4% were perimenopausal (≥40 and <180 days since last menstrual period for at least one observation or changed menopausal status during follow-up). Average ages were 56 years for postmenopausal women and 41

Table 1 Characteristics of the Gilda Radner Ovarian Cancer cohort, 1991 to 1997

	n	%	Mean	SD	Min.[a]	Max.
Age at enrollment (yr)						
All women	1257	100.0	45.5	9.8	21	80
Premenopausal	790	62.9	40.5	5.8		
Perimenopausal	105	8.4	46.8	5.2		
Postmenopausal	346	27.5	56.4	9.2		
Unknown	16	1.3	47.1	8.0		
No. of screenings	5358	100.0	4.3	2.9	1	14
Days between screenings			237	176	1	1953

[a] Min., minimum; Max., maximum.

Table 2 Select tumor marker levels (units/ml) by menopausal status

Marker	Women	*n*	Mean	SD	Median	Max.[a]
CA125	All	5266	24.77	74.34	16	3400
	Premenopausal	3487	29.04	80.63	19	3400
	Perimenopausal	133	16.03	12.88	14	124
	Postmenopausal	1612	15.18	48.10	12	1900
HER-2/*neu*	All	5242	12.96	37.29	11	1680
	Premenopausal	3472	13.18	45.48	11	1680
	Perimenopausal	133	12.73	5.29	12	41
	Postmenopausal	1603	12.52	7.98	12	152
UGP	All	4362	1.63	23.50	0	1351
	Premenopausal	2914	1.26	28.66	0	1351
	Perimenopausal	113	2.56	4.95	0.8	38
	Postmenopausal	1307	2.41	3.31	1.4	41
LASA	All	5261	15.48	3.96	15	59.1
	Premenopausal	3166	15.31	3.64	15	41.0
	Perimenopausal	235	15.22	3.85	15	28.5
	Postmenopausal	1606	15.94	4.41	16	59.1
DM/70K	All	5262	5.16	10.89	0	253
	Premenopausal	3166	3.96	9.00	0	158
	Perimenopausal	235	7.66	13.06	0	119
	Postmenopausal	1606	7.25	13.35	0	253

[a] Max., maximum.

years for premenopausal women. Descriptive statistics for all five tumor markers by menopausal status are shown in Table 2.

Tumor Marker Behavior. Tumor marker levels were compared by menopausal status and evaluated for heterogeneity among women. CA125 and HER-2/*neu* mean levels were significantly lower among postmenopausal women than among premenopausal women, whereas LASA, DM/70K, and UGP were significantly higher among postmenopausal women (Kruskal-Wallis test, $P = 0.0001$ for all comparisons). CA125, DM/70K, UGP, and HER-2/*neu* distributions were highly skewed, and for subsequent analysis were converted to an approximately normal distribution using the natural logarithm transformation. Two percent of postmenopausal women and 15% of premenopausal women had first-screen CA125 values exceeding 35 units/ml, a frequently used cutoff level. Overall, CA125 levels were somewhat higher than observed in other cohorts, but the variance decomposition (the focus of this analysis) was similar to other cohort data on the logarithmic scale, which appears more appropriate for longitudinal modeling (12).

ICC was estimated separately for each tumor marker. High ICC implies that a large proportion of data variability occurs among individual mean levels, and thus is suggestive of true heterogeneity among women. When ICC is substantial in magnitude, one can expect substantial improvement when using information from longitudinal screening compared with a single threshold rule (12). Estimates of ICC were approximately 0.6 for log CA125 and log UGP, 0.4 for LASA, 0.3 for log HER-2/*neu*, and 0.2 for log DM/70K (Table 3). These estimates suggest considerable heterogeneity in individual tumor marker means, particularly for log CA125 and log UGP.

Table 3 Tumor marker variability and intraclass correlation

Tumor marker	No. of measurements	No. of women	SD[a]			ICC[b]
			Overall	σ	τ	
log CA125	5266	1253	0.758	0.483	0.584	0.594
log HER-2/*neu*	5239	1250	0.483	0.412	0.252	0.272
log UGP	1885	827	1.046	0.698	0.779	0.555
LASA	5261	1253	3.962	3.113	2.451	0.383
log DM/70K	2092	798	0.975	0.880	0.421	0.187

[a] σ, within-person standard deviation; τ, between-person standard deviation.
[b] Intraclass correlation.

For all markers, individual means were significantly heterogeneous, regardless of menopausal status (overall F test, $P < 0.001$). Within-person variances were also significantly heterogeneous for all five markers (Bartlett's test, $P < 0.001$). These findings corroborated the ICC evidence for substantial heterogeneity in tumor marker patterns in this cohort.

The five tumor markers we studied behaved approximately independently as indicated by small pairwise correlations. Pearson correlation coefficients based on the total cohort data ranged from −0.13 to 0.26 (Table 4). When examined by menopausal status, the magnitudes of these correlations were not substantially changed (data not shown).

Development of individual-specific screening rules relies on the estimation of tumor marker parameters for each individual undergoing screening. To illustrate this in the Gilda Radner cohort, the hierarchical Bayes model was used to estimate the mean, variance, and correlation of log CA125 and log HER-2/*neu* for all postmenopausal women who had at least three measurements of these markers ($n = 323$). Table 5 shows these estimates: the cohort mean (μ_0) and between-person SD (τ) for log CA125 and log HER-2/*neu*; and the between-person ρ, an estimate of correlation between individual-specific log CA125 and log HER-2/*neu* means. For two representative individuals, estimates of within-person mean (μ_i), SD (σ_i), and correlation (ρ_i) of log CA125 and log HER-2/*neu* are presented also (Table 5). Individual-specific parameters estimated in this way may be used to identify more optimally any tumor marker increase from an individual's naturally occurring mean level. For example, in Table 5, person 1 has a log CA125 mean of 2.63 and SD of 0.34; person 2 has a mean and SD of 2.85 and 0.43, respectively. If these values were known to be reliable, an upper bound for normal could be fixed for person 1 at $2.63 + 2 \times 0.34 = 3.31$, and for person 2 at $2.85 + 2 \times 0.43 = 3.71$. Because a standard cutoff is log (35) = 3.56, this implies that the first woman should have a lower cutoff than the standard, and the second should have a higher cutoff than the standard. The statistical details are complicated for how to implement this rule when individual-level tumor marker data are incomplete or uncertain. Skates *et al.* (12) have proposed an algorithm for CA125 that can be adapted to other single markers, but specific algorithms that can capture information from multiple markers in this manner have yet to be proposed.

Discussion

Screening protocols using serum tumor markers generally have relied on a fixed cutoff level to determine abnormal values.

Table 4 Correlations (ρ) between tumor markers

	CA125 (log CA125)	HER-2/*neu* (log H2N)	UGP (log UGP)	LASA (LASA)	DM/70K (log DM/70K)
CA125	1.0000	0.0057	−0.0004	0.0693	0.0793
(log CA125)	(1.0000)	(0.1720)	(−0.1315)	(0.2161)	(0.0697)
HER-2/*neu*		1.0000	−0.0023	0.0236	−0.0038
(log HER-2/*neu*)		(1.0000)	(0.1081)	(0.2540)	(0.0789)
UGP			1.0000	0.0136	0.0212
(log UGP)			(1.0000)	(0.0756)	(0.1339)
LASA				1.0000	0.2790
(LASA)				(1.0000)	(0.2599)
DM/70K					1.0000
(log DM/70K)					(1.0000)

Table 5 CA125 and HER-2/*neu*: bivariate estimates for all postmenopausal women and individual-specific estimates for two representative postmenopausal women

	Point estimate	95% CI[a]
Between-person[b]		
Log CA125 mean (μ_0)	2.90	2.84–2.96
Log CA125 SD (τ)	0.49	0.45–0.54
Log HER-2/*neu* mean (μ_0)	2.40	2.37–2.42
Log HER-2/*neu* SD (τ)	0.18	0.02–0.21
Between-person correlation (ρ)	0.12	−0.05–0.28
Within-person[c]		
Person 1		
Log CA125 mean (μ_i)	2.63	2.38–2.89
Log CA125 SD (σ_i)	0.34	0.22–0.54
Log HER-2/*neu* mean (μ_i)	2.29	2.10–2.48
Log HER-2/*neu* SD (σ_i)	0.29	0.19–0.42
Within-person correlation (ρ_i)	0.03	−0.09–0.20
Person 2		
Log CA125 mean (μ_i)	2.85	2.54–3.16
Log CA125 SD (σ_i)	0.43	0.29–0.66
Log HER-2/*neu* mean (μ_i)	2.46	2.26–2.66
Log HER-2/*neu* SD (σ_i)	0.34	0.24–0.47
Within-person correlation (ρ_i)	0.03	−0.07–0.21

[a] CI, confidence interval.
[b] Between-person estimates are based on all postmenopausal women having at least three observations ($n = 323$).
[c] Within-person estimates are based on two representative individuals, each having seven observations.

Women at risk for ovarian cancer whose CA125 levels exceed 35 units/ml, for example, will undergo more vigilant monitoring with imaging modalities or further clinical evaluation. Such protocols have not adequately used information on the heterogeneity of tumor markers in a screening population. If tumor marker patterns are naturally heterogeneous among women, the use of a fixed cutoff level is a crude, suboptimal method for detecting increases from each individual's baseline. The goals of this analysis were to characterize the normal behavior of five candidate tumor markers in a large ovarian cancer screening cohort and to highlight the implications for screening.

The CA125 mean level was significantly lower and the UGP mean level significantly higher among postmenopausal compared with premenopausal women. These findings have been reported previously (15, 21, 22). We also observed that DM/70K and LASA mean levels were higher, and the HER-2/*neu* mean level lower, among postmenopausal compared with premenopausal women. For LASA and HER-2/*neu* these differences were slight, but all comparisons were statistically significant.

Individual-specific tumor marker characteristics were also estimated for this cohort of 1257 women. These women had very heterogeneous patterns of all five markers studied, particularly CA125. The overall CA125 mean was 25 units/ml with wide 95% reference centiles (6–75 units/ml), indicating substantial heterogeneity of individual mean levels. The average CA125 SD was 12 units/ml, also with wide 95% reference centiles (0.7–56 units/ml), suggesting substantial heterogeneity of individual-specific variability. Fifteen percent of premenopausal women and 2% of postmenopausal women had a first-screen CA125 level exceeding 35 units/ml, and 68% and 57% of these women, respectively, had a recurrent elevation on the second screen without development of ovarian cancer. ICC, the best summary indicator of heterogeneity among individuals, was nearly 0.6 for log CA125 and log UGP, and less than 0.4 for LASA, log DM/70K, and log HER-2/*neu*. These findings suggest that tumor marker patterns are substantially heterogeneous even among healthy, cancer-free women.

Approximate independence of the tumor markers that we studied gives a clinically useful result. The combined false-positive rate from screening with multiple markers is well estimated by the sum of individual false-positive rates of the markers, provided that the specificity of markers is reasonably high (as is usually the case for reasonable screening candidates). For instance, in postmenopausal women that we studied, a CA125 cutoff level of 35 units/ml gave a specificity of 0.98 (the probability of a negative test in the absence of disease), and the HER-2/*neu* standard cutoff level of 20 units/ml gave a specificity of 0.95. Because these markers were approximately independent, the combined specificity of both tests used jointly is $0.98 \times 0.95 = 0.931$, and the probability of at least one false-positive test is $1 - 0.931 = 0.069$. The latter is very closely estimated by the sum of individual false-positive rates (1 − specificity) for these markers, $0.02 + 0.05 = 0.07$.

The heterogeneity of tumor marker patterns observed among women in this cohort underscores the need for incorporating individual-specific decision rules in screening protocols. To date, most diagnostic tests using tumor markers have not accounted for a woman's screening history in the evaluation of tumor marker levels. In screening, a fixed cutoff level is suboptimal to a degree that depends strongly on the ICC. It is because of this phenomenon that the algorithm by Skates *et al.* (12) performed better than the conventional use of CA125. Using serial measurements of tumor markers, individual-specific screening rules may be developed that use all available information on the individual level as well as on the screening population level. This approach to screening, extended to multiple markers, will require information on ICCs of markers such as are presented in this study.

References

1. American Cancer Society. Cancer Facts and Figures: 2000. New York: American Cancer Society, 2000.

2. Bast, R. C., Jr., Siegal, F. P., Runowicz, C., Klug, T. L., Zurawski, V. R., Jr., Schonholz, D., Cohen, C. J., and Knapp, R. C. Elevation of serum CA125 prior to diagnosis of an epithelial ovarian carcinoma. Gynecol. Oncol., *22:* 115–120, 1985.

3. Mann, W. J., Patsner, B., Cohen, H., and Loesch, M. Pre-operative serum CA125 levels in patients with surgical stage I invasive ovarian adenocarcinoma. J. Natl. Cancer Inst., *80:* 208–209, 1989.

4. Jacobs, I. J., Skates, S., Davies, A. P., Woolas, R. P., Jeyerajah, A., Weidemann, P., Sibley, K., and Oram, D. H. Risk of diagnosis of ovarian cancer after raised serum CA 125 concentration: a prospective cohort study. Br. Med. J., *313:* 1355–1358, 1996.

5. Nam, J-H., Cole, L. A., Chambers, J. T., and Schwartz, P. E. Urinary gonadotropin fragment, a new tumor marker. I. Assay development and cancer specificity. Gynecol. Oncol., *36:* 383–390, 1990.

6. Dnistrian, A. M., and Schwartz, M. K. Plasma lipid-bound sialic acid and carcinoembyonic antigen in cancer patients. Clin. Chem., *27:* 1737–1739, 1981.

7. Karlan, B. Y., Raffel, L. J., Crvenkovic, G., Smrt, C., Chen, M. D., Lopez, E., Walla, C. A., Garber, C., Cane, P., and Sarti, D. A. A multidisciplinary approach to the early detection of ovarian carcinoma: rationale, protocol design, and early results. Am. J. Obstet. Gynecol., *169:* 494–501, 1993.

8. Knauf, S., Anderson, D. J., Knapp, R. C., and Bast, R. C., Jr. A study of the NB/70K and CA125 monoclonal antibody radioimmunoassays for measuring serum antigen levels in ovarian cancer patients. Am. J. Obstet. Gynecol. *152:* 911–913, 1985.

9. Slamon, D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., and Ullrich, A. Studies of the HER-2/*neu* proto-oncogene in human breast and ovarian cancer. Science (Washington DC), *244:* 707–712, 1989.

10. Karlan, B. Y. Screening for ovarian cancer: what are the optimal surrogate endpoints for clinical trials? J. Cell. Biochem. Suppl., *23:* 227s–232s, 1995.

11. Cirisano, F. D., and Karlan, B. Y. The role of the HER-2/*neu* oncogene in gynecologic cancers. J. Soc. Gynecol. Investig., *3:* 99–105, 1996.

12. Skates, S. J., Xu, F. J., Yu, Y. H., Sjovall, K., Einhorn, N., Chang, Y., Bast, R. C., Jr., and Knapp, R. C. Toward an optimal algorithm for ovarian cancer screening with longitudinal tumor markers. Cancer (Phila.), *76* (Suppl.)*:* 2004s–2010s, 1995.

13. Woolas, R. P., Conaway, M. R., Xu, F., Jacobs, I. J., Yu, Y., Daly, L., Davies, A. P., O'Briant, K., Berchuck, A., and Soper, J. T. Combinations of multiple serum markers are superior to individual assays for discriminating malignant from benign pelvic masses. Gynecol. Oncol., *59:* 111–116, 1995.

14. Bast, R. C., Jr., Boyer, C. M., Xu, F. J., Wiener, J., Dabel, R., Woolas, R., Jacobs, I., and Berchuck, A. Molecular approaches to prevention and detection of epithelial ovarian cancer. J. Cell. Biochem. Suppl., *23:* 219s–222s, 1995.

15. Cane, P., Azen, C., Lopez, E., Platt, L. D., and Karlan, B. Y. Tumor marker trends in asymptomatic women at risk for ovarian cancer: relevance for ovarian cancer screening. Gynecol. Oncol., *57:* 240–245, 1995.

16. Karlan, B. Y., Baldwin, R. L., Lopez-Luevanos, E., Raffel, L. J., Barbuto, D., Narod, S., and Platt, L. D. Peritoneal serous papillary carcinoma, a phenotypic variant familial ovarian cancer: implications for ovarian cancer screening. Am. J. Obstet. Gynecol., *180:* 917–925, 1999.

17. Casella, G. An introduction to empirical Bayesian data analysis. American Statistician, *39:* 83–87, 1985.

18. Morris, C. N. Parametric empirical Bayes inference: theory and applications. J. Am. Stat. Assoc., *78:* 47–65, 1983.

19. Efron, B., and Morris, C. Stein's paradox in statistics. Sci. Am., *236:* 119–127, 1977.

20. DuMouchel, W. Documentation for Hierarchical Bayes Linear Model Program. New York: Columbia University Department of Medical Informatics, 1995.

21. Schwarz-Roeger, U., Petzoldt, B., Waldschmidt, R., Walker, R. P., Bauknecht, T., and Kiechle, M. UGP: a tumor marker of gynecologic and breast malignancies? Specificity and sensitivity in pretherapeutic patients and the influence of hormonal substitution on the expression of UGP. Anticancer Res., *17:* 3041–3045, 1997.

22. Walker, R., Crebbin, V., Stern, J., Scudder, S., and Schwartz, P. Urinary gonadotropin peptide (UGP) as a marker of gynecologic malignancies. Anticancer Res., *14:* 1703–1709, 1994.

Ovarian Cancer Tumor Marker Behavior in Asymptomatic Healthy Women: Implications for Screening

Casey Crump, Martin W. McIntosh, Nicole Urban, et al.

Cancer Epidemiol Biomarkers Prev 2000;9:1107-1111.