

Chemistry. In the article “Virtual combinatorial libraries: Dynamic generation of molecular and supramolecular diversity by self-assembly,” by Ivan Huc and Jean-Marie Lehn, which appeared in number 6, March 18, 1997, of *Proc. Natl. Acad. Sci. USA* (**94**, 2106–2110), the following changes should be noted. On page 2106, the first line in the second paragraph should read “Dynamic combinatorial chemistry is a conceptually different. . . .” On page 2110, the fourth line in the second paragraph should read “. . . the assembly-forming connections scheme of brain function (34). . . .”

Medical Sciences. In the article “The CAG repeat within the androgen receptor gene and its relationship to prostate cancer,” by Edward Giovannucci, Meir J. Stampfer, Krishna Krithivas, Myles Brown, Adam Brufsky, James Talcott, Charles H. Hennekens, and Philip W. Kantoff, which appeared in number 7, April 1, 1997, of *Proc. Natl. Acad. Sci. USA* (**94**, 3320–3323), the authors wish to point out the omission of coauthor Douglas Dahl from their paper. The corrected author line should read as follows. Edward Giovannucci, Meir J. Stampfer, Krishna Krithivas, Myles Brown, Douglas Dahl, Adam Brufsky, James Talcott, Charles H. Hennekens, and Philip W. Kantoff.

Dr. Dahl’s affiliation is the Division of Medical Oncology, Dana–Farber Cancer Institute, and the Division of Urology, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA 02115.

The CAG repeat within the androgen receptor gene and its relationship to prostate cancer

EDWARD GIOVANNUCCI*[†], MEIR J. STAMPFER*[‡], KRISHNA KRITHIVAS[§], MYLES BROWN[¶], ADAM BRUFKY^{§¶}, JAMES TALCOTT[§], CHARLES H. HENNEKENS^{‡||}, AND PHILIP W. KANTOFF[§]

*Channing Laboratory, Brigham and Woman's Hospital and Harvard Medical School, and Department of Nutrition, Harvard School of Public Health, Boston, MA 02115; [‡]Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115; Divisions of [§]Medical Oncology and [¶]Neoplastic Disease Mechanisms, Dana–Farber Cancer Institute, Harvard Medical School, Boston, MA 02115; and Division of Preventative Medicine, Brigham and Woman's Hospital and Harvard Medical School, Boston, MA 02115

Communicated by Arthur B. Pardee, Dana–Farber Cancer Institute, Boston, MA, January 21, 1997 (received for review December 20, 1996)

ABSTRACT The length of a polymorphic CAG repeat sequence, occurring in the androgen receptor gene, is inversely correlated with transcriptional activity by the androgen receptor. Because heightened androgenic stimulation may increase risk of prostate cancer development and progression, we examined whether shorter CAG repeats in the androgen receptor gene are related to higher risk of prostate cancer. We conducted a nested case-control study of 587 newly diagnosed cases of prostate cancer detected between 1982 and 1995, and 588 controls without prostate cancer, within the Physician's Health Study. An association existed between fewer androgen receptor gene CAG repeats and higher risk of total prostate cancer [relative risk (RR) = 1.52; 95% confidence interval (CI) = 0.92–2.49; *P* trend = 0.04; for men with CAG repeat lengths ≤ 18 relative to ≥ 26 repeats]. In particular, a shorter CAG repeat sequence was associated with cancers characterized by extraprostatic extension or distant metastases (stage C or D) or high histologic grade (RR = 2.14; CI = 1.14–4.01; *P* trend = 0.001). This association was observed individually both for high stage (RR = 2.23) and high grade prostate cancer (RR = 1.89). Men with shorter repeats were at particularly high risk for distant metastatic and fatal prostate cancer. Variability in the CAG repeat length was not associated with low grade or low stage disease. These results demonstrate that a shorter CAG repeat sequence in the androgen receptor gene predicts higher grade and advanced stage of prostate cancer at diagnosis, and metastasis and mortality from the disease. The clinical implications of these results should be evaluated further.

Cell division in the prostate gland is controlled by the interaction of testosterone and dihydrotestosterone with the androgen receptor (AR) (1). The AR gene contains a polymorphic CAG repeat sequence, which ranges normally from about 8 to 31 repeats and averages about 20 (2). This CAG repeat sequence, present in exon 1, encodes for a polyglutamine chain in the region of the AR connected with DNA transcription. *In vitro*, the length of the polyglutamine chain correlates inversely with transcriptional activity by the AR (3, 4). Men who possess exceptionally long CAG repeat lengths experience clinical androgen insensitivity (5–7), presumably related to ARs possessing reduced transcriptional activity.

Prostate carcinogenesis is dependent on androgens (1). Because shorter CAG repeat lengths are associated with high transcriptional activity of the AR, Irvine *et al.* (8) have proposed that men with shorter repeat lengths will be at higher

risk of prostate cancer. Some indirect evidence is consistent with this hypothesis. African Americans, who have generally shorter CAG repeat lengths in the AR (9), have a higher incidence and mortality rate from prostate cancer. Moreover, the AR is located on the X chromosome, and consistent with an X-linked genetic component for prostate cancer is that history of the disease in a brother carries greater risk than paternal history (10–15). These intriguing observations led us to assess directly whether polymorphism in CAG repeat length in the AR is related to prostate cancer development and progression in case-control study conducted in the Physician's Health Study.

METHODS

Study Population. The cases and controls were selected from the Physician's Health Study, a randomized double-blind trial of aspirin and beta-carotene among 22,071 U.S. male physicians, aged 40–84 years in 1982 (16). The cohort is predominantly caucasian (over 95%). Men were excluded if they reported a prior history of myocardial infarction, stroke, transient ischemic attacks, unstable angina, cancer (except for nonmelanoma skin cancer), current renal or liver disease, peptic ulcer or gout, contraindication to use of aspirin, or current use of aspirin, other platelet-active agents, or vitamin A supplements.

Study participants completed two mailed questionnaires before randomization in 1982, and additional questionnaires at 6 months, 12 months, and annually thereafter. Before randomization we sent blood kits to all participants with instructions to have their blood drawn into vacutainer tubes containing EDTA, to centrifuge them, and to return the specimens (by overnight prepaid courier) in polypropylene cryopreservation vials. The kit included a cold pack to keep the specimens cool until receipt the following morning, when they were aliquoted and stored at -82°C . We received specimens from 14,916 (68%) of the randomized physicians; these men form our analytic cohort.

Selection of Cases and Controls. After 13 years of follow-up, over 99% of the men were still responding to follow-up questionnaires, and vital status was ascertained for 100%. When a participant reported a diagnosis of cancer on the follow-up questionnaires, we requested medical records, including pathology reports, that were reviewed by study physicians from the End Points Committee. By 1995, we confirmed 591 cases of prostate cancer among the 14,916 who had provided blood. For each case, we selected one control who had provided blood, had not had a previous prostatectomy, and had not reported a diagnosis of prostate cancer at the time diagnosis was reported by the case. Controls were also matched

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA
0027-8424/97/943320-4\$2.00/0
PNAS is available online at <http://www.pnas.org>.

Abbreviations: AR, androgen receptor; RR, relative risk; CI, confidence interval.

on smoking status and age within 1 year, except for several elderly cases for whom age had to be matched within 2 years.

Ascertainment of Prostate Cancer. A study physician, unaware of assay results, reviewed medical records for each case to determine stage at diagnosis, tumor grade, and Gleason score. Stage was recorded according to the modified Whitmore–Jewett classification scheme (17). If multiple tissue samples were examined, the highest reported grade and Gleason score were recorded. Cases without pathological staging were classified as indeterminate stage unless there was clinical evidence of distant metastases. High grade/stage cases were defined as those diagnosed at stage C or D (extraprostatic) plus those diagnosed at stage A or B or indeterminate with either poor histologic differentiation or Gleason score 7 or greater. Cases with clinical stage A or B prostate cancer or with no pathological staging and moderate or better histologic grade were classified as low grade/stage. Fatal prostate cancer consisted of newly diagnosed cases that were fatal by the end of follow-up.

Analysis for CAG Repeat Length in the AR. Since the AR gene is X-linked, only one copy of the gene exists in males. The CAG repeat region resides in the first exon of the gene. We established a system to rapidly analyze the CAG repeat sequence length in a large number of samples. Five hundred microliters of whole blood was thawed from cases and controls, and DNA was extracted utilizing the Qiagen QIAamp Blood Kit (Qiagen, Chatsworth, CA). We constructed a set of oligonucleotide primers that flank the CAG repeat (5'-TCCAGAATCTGTTCCAGAGCGTGC-3' and 5'-GCTGTGAAGGTTGCTGTTCCCTCAT-3'). The DNA was amplified using these primers by PCR to produce fragments of the N-terminal domain of the AR. Primers were fluorescently labeled. The length of these fragments varied only by the number of CAG repeats. For rapid and accurate assessment of fragment length, the DNA fragments were run on a 6% denaturing polyacrylamide gel by automated fluorescence detection (Genescan; Applied Biosystems) in the Dana-Farber Cancer Institute Molecular Biology Core Facility. Using a series of sequenced PCR products of varying size, fluorescently labeled DNA markers were used to create a standard curve of peak arrival time, allowing us to calculate the length of an unknown PCR product. Resolution to one 1-bp length using this system was confirmed with direct DNA sequencing. The assays were conducted by laboratory personnel blinded to case-control status. We utilized split samples to ensure quality control. We were able to amplify the DNA for 587 of the 591 cases and 588 of the 591 controls (>99%).

Data Analysis. We conducted analyses to determine whether the AR CAG repeat sequence length was related to prostate cancer risk. We used unconditional logistic regression (18), controlling for age and smoking, the matching variables, to compute the relative risk or RR (estimated by the odds ratio) of prostate cancer, and 95% confidence intervals (CI). Results using conditional logistic regression for matched analyses were very similar to those using unconditional logistic regression and controlling for the matching variable. By using unconditional logistic regression, we were able to use all of the controls in the subgroup analysis of cases, such as metastatic or fatal cases, thus enhancing statistical power. This multivariate analysis allows us to control for potentially confounding effects of other variables, including age, alcohol intake, physical activity, multivitamin use, body mass index, and aspirin use. All reported *P* values are two-sided.

Because AR transcriptional activity decreases linearly across the entire CAG spectrum (3, 4) we analyzed CAG repeats as a continuous variable in logistic regression models. This approach assumes that each one-unit increment in CAG repeat length is related to a constant proportional change in the relative risk. We present the calculated RR for prostate cancer for a six-CAG decrement, which is equivalent to the

difference between the median CAG repeat length in the high and low tertiles. This RR is calculated by multiplying by six the beta coefficient from a logistic regression model where CAG repeat length is treated as a continuous variable, and exponentiating the product.

To complement the continuous analysis, we classified men into categories of CAG repeats to observe if nonmonotonic relationships, such as a threshold effect, were apparent. The categorization (ranging from ≤ 18 to ≥ 26 repeats) was based on approximating relatively equal numbers in the categories, although the numbers fluctuated somewhat because of the uneven distribution of CAG repeats. We used the category of ≥ 26 repeats as the reference or comparison group, and examined if the relative risk increased incrementally in groups with progressively shorter repeat lengths.

In addition to total prostate cancer, we conducted analyses of tumors that possessed a more aggressive phenotype as determined by histology (tumor grade or Gleason score), tumor stage, and fatality. Our rationale was that prostate cancer may consist of subtypes that vary in aggressive behavior or that factors that influence initiation may differ from those that influence progression. This heterogeneity in aggressive behavior or progression rate among prostate cancers is suggested by the fact that rates of the aggressive type of prostate cancer vary substantially among populations whereas frequency of the latent, noninfiltrative type of cancer does not vary (19).

RESULTS

The number of CAG repeats in the AR ranged from 12 to 35 among cases and from 6 to 39 among controls. Among the controls, the mode of the distribution occurred at 21 repeats (17% of men), $\approx 10\%$ of the men fell in each category of 19, 20, 22, 23, 24, and 25 repeats. Abrupt drops in the frequencies of men occurred above 25 repeats and below 19 repeats. The normal range of repeats, 6–39, was wider than previously reported (3). Among 587 total cases and 269 high grade/high stage cases, no apparent correlation in AR CAG repeat length was noted with age ($P > 0.3$).

We observed a progressive increase in risk of prostate cancer associated with decreasing CAG repeat length (Table 1). This relationship was noted for cancers characterized as high grade or advanced stage, but not for low grade or low stage cancers. This association was observed individually both for high stage (RR = 2.23, CI = 1.08–4.61, comparing men with repeat lengths ≤ 18 to those ≥ 26 ; P trend = 0.002) and high grade cancers (RR = 1.89, CI = 0.96–3.75; P trend = 0.007). AR CAG repeat length was not correlated with any cofactor considered (year of diagnosis, alcohol intake, physical activity, multivitamin use, body mass index, and aspirin use); hence, the results were unaltered when these were included as covariates in the logistic regression models.

For the categorical analysis, statistically significant results were noted only when comparing the extreme categories of CAG repeats, although the general pattern appeared consistent with a moderate progressive increase in risk of aggressive prostate cancer associated with decreasing CAG repeat length. No obvious threshold or nonlinear patterns were apparent. We thus analyzed CAG repeats as a continuous variable in logistic models, and present an odds ratio to estimate relative risk for a decrement of six CAG repeats (Table 2). As for the categorical analysis, this analysis demonstrated that men possessing fewer CAG repeats are at higher risk of advanced stage, high grade, distant metastatic, and fatal prostate cancer. For example, based on our results, a man with a CAG repeat length of 19 is at ≈ 2.44 higher risk of developing distant metastatic prostate cancer as a man with a CAG repeat length of 25.

We had initially observed these relationships in 367 cases mostly diagnosed by 1991 before the widespread use of pros-

Table 1. Odds ratio of prostate cancer by length of CAG repeat length of the AR gene among men in the Physicians' Health Study (1982–1995)

	CAG length							P value (trend)
	≥26	24–25	22–23	21	20	19	≤18	
Controls	72	115	119	101	65	61	55	
Total prostate cancer (n)	60	98	116	113	69	62	69	
OR*	1.0	1.02	1.17	1.35	1.28	1.22	1.52	0.04
95% CI		(0.66–1.58)	(0.76–1.80)	(0.87–2.09)	(0.79–2.08)	(0.75–2.00)	(0.92–2.49)	
High grade/stage prostate cancer (n)†	24	37	47	55	38	30	38	
OR*	1.0	0.99	1.20	1.65	1.79	1.49	2.14	0.001
95% CI		(0.55–1.80)	(0.68–2.14)	(0.93–2.92)	(0.97–3.32)	(0.79–2.82)	(1.14–4.01)	
Low grade/stage prostate cancer (n)	35	58	69	58	30	31	28	
OR*	1.0	1.03	1.20	1.18	0.94	1.06	1.03	0.86
95% CI		(0.61–1.99)	(0.72–1.99)	(0.70–1.99)	(0.52–1.71)	(0.58–1.92)	(0.56–1.91)	

*Odds ratio (OR) and 95% CI for various categories of CAG repeats relative to men with repeat size ≤19. OR controlled for age (in 5-year age categories) and smoking (past, current) by unconditional logistic regression.

†Tumors with Gleason grade ≥7 or high grade, or those with extension outside the prostate gland (stage C or D).

tatic-specific antigen (PSA) for screening. We subsequently confirmed this association in 220 new cases diagnosed after March 1992, during the era of prevalent use of PSA for screening. The combined 587 cases comprise the cases described in this report. The relative risks were very similar in the initial analysis [for high stage/grade lesions, RR (for a CAG decrement of 6) = 1.52 (95% CI = 1.04–2.22; $P = 0.03$), and RR = 1.92 (CI = 1.10–3.45; $P = 0.02$, for cases during the subsequent time period)]. No appreciable association was observed for low grade/stage cancers during either time period.

DISCUSSION

Various lines of evidence suggest that the occurrence and progression of malignancies of this gland are influenced by androgen-stimulated cell division. Cancers of the prostate gland are generally sensitive to androgens and often regress initially when androgen stimulation is withdrawn (1). Malignancies of the prostate occur rarely in castrated men (20), and the prolonged administration of high levels of testosterone has induced prostate cancer in rats (21, 22). While abnormally low levels of androgens are associated with low risk of the disease and exceptionally high levels induce cancer in animals, whether androgenic stimulation within the normal range in males is associated with moderate differences in risk is unsettled, though supported by recent evidence (23).

Table 2. Odds ratio of prostate cancer for a CAG microsatellite repeat length decrement of six in the androgen receptor gene among men in the Physicians' Health Study

Prostate Cancer	Cases	Odds ratio*		P
		(six decrement in CAG)	95% CI	
Total	587	1.28	(1.01–1.61)	0.04
High grade/stage†	269	1.64	(1.22–2.22)	0.001
Low grade/stage	309	1.02	(0.77–1.37)	0.86
High grade	210	1.59	(1.14–2.22)	0.007
Advanced stage	180	1.75	(1.23–2.50)	0.002
Metastatic (distant)	56	2.44	(1.32–4.55)	0.004
Fatal	43	2.08	(1.05–4.00)	0.04

*Odds ratio is calculated by modeling CAG as a continuous variable in an unconditional logistic model and computing the odds ratio for a six CAG decrement (difference from median of low to median of high tertile of CAG repeat length).

†Includes tumors with Gleason grade ≥7 or high grade or advanced stage (C or D).

The action of androgens is ultimately mediated through the AR. In transfection experiments, longer AR polyglutamine repeat lengths encoded by CAG repeats are associated with lower transcriptional activity. Two laboratories (3, 4) have independently established that this relationship is length-dependent, and occurs even within the normal range of CAG repeats. Binding of androgens to the AR occurs in a different region of the AR; this polymorphism in polyglutamine length does not appear to influence binding of androgens to the AR. Abnormally high CAG repeat length sequences (≥40), which through an unknown mechanism cause spinobulbar muscular atrophy or Kennedy syndrome, is associated with clinically overt androgen insensitivity in men (5, 7). Based on a small sample ($n = 16$), women with normal testosterone levels but with idiopathic hirsutism exhibited an inverse correlation between degree of hirsutism and CAG repeat size within the normal range ($r = -0.60$, $P = 0.01$) (24).

Given clear evidence of clinical androgen insensitivity with long CAG repeat lengths and the linear gradient between CAG repeat length and AR transcriptional activation *in vitro*, a reasonable supposition is that variation within the normal range is associated with differences in transcriptional activity, albeit modest, *in vivo*. Based on the assumption that androgens are critical to prostate cancer development or progression, Coetzee and Ross (9) had hypothesized that variation in transcriptional activity by the AR, related to polymorphic CAG repeats, influences prostate carcinogenesis. We tested this hypothesis in a large, prospective study, and found that shorter CAG repeat lengths of the AR was closely related to an aggressive phenotype of prostate cancer, as defined by high histological grade, extension through the prostate gland, presence of distant metastasis at diagnosis, and mortality from the disease. A highly significant association occurred independently for both tumor grade and stage, increased in magnitude with degree of aggressive behavior, such as distant metastases and mortality, and occurred consistently over time in this cohort, arguing strongly against a chance finding.

Based on the *in vitro* study by Kazemi-Esfarjani *et al.* (4) we estimated approximately a 12% differential in transcriptional activation for a decrement of six CAG repeats. Although the magnitude of the effect of the AR polyglutamine length and transcriptional activation function *in vitro* may appear relatively modest, these differences over a lifetime might have a substantial impact. Using a mathematical model that assumes that prostate cancer risk is directly proportional to cumulative mitotic activity, Ross (25) has estimated that a 13% difference in androgen-stimulated mitotic activity would result in a 2.8-fold difference in prostate cancer incidence. For a decre-

ment of six CAG repeats or about a 12% predicted difference in transcriptional activation, our data estimated relative risks of 2.44 for metastatic disease and 2.08 for fatal disease, which are consistent with the magnitude predicted from the model.

A study of 109 men referred to a medical oncology practice at Memorial Sloan-Kettering Cancer Center found that shorter repeat lengths were significantly correlated with younger age of diagnosis of prostate cancer (26). Most known germ-line mutations that confer higher risk of cancer (e.g., BRCA1 in breast cancer, mismatch repair genes in colon cancer) are characterized by early age of disease onset, high population attributable risk at young ages, but a relatively low population attributable risk due to the sharply increasing incidence of "sporadic" cancers that occurs with advancing age. In contrast, our findings are consistent with a moderate gradient of risk occurring across the spectrum of CAG repeats and across age groups. Because this polymorphism influences the progression of "sporadic" cancers, including those occurring at older ages, the population attributable risk may be quite high. It is critical for future studies to examine the interactions among serum testosterone level, which decrease with age, AR CAG repeat length, and age. Clearly, a substantially larger sample size than the current study would be required for sufficient power to examine these interactions.

African-American men have about a 2-fold higher rate of metastatic prostate disease and mortality, and have larger tumor volumes, even when they have equal access to health care as whites (27, 28). Although the similar access to care does not assure equivalent utilization, it would appear that black men are prone to a more aggressive tumor biology. Because black men tend to have considerably shorter AR CAG repeats than white men, we speculate that CAG repeat length may be a factor contributing to an increased risk of metastatic and fatal prostate cancer in blacks. However, our population was predominantly caucasian (>95%), so this hypothesis needs to be studied directly in an African-American population.

Our results provide strong evidence that the variability in the CAG repeat length, probably through mediation of transcriptional activation of the AR, is associated with the risk of developing prostate cancer and in particular aggressive prostate cancer. This finding may help explain the higher rate of prostate cancer mortality among African-Americans, as well as their tendency to be diagnosed with more extensive disease, and the apparent X-linked component to prostate cancer risk. Although our results are consistent with a substantial effect of CAG repeat length, further study is required before the appropriate clinical relevance can be established.

We would like to acknowledge Paul Morrison for his assistance in establishing the assay. This work was supported by National Institutes of Health Cancer Research Grants HL-26490, HL-34595, CA-34944, CA-42182, and CA-58684, a CaPCURE grant, and by the Dana-Farber Cancer Institute Prostate Cancer Research Fund.

1. Coffey, D. S. (1979) *Prostate Cancer*, UICC Technical Report Series (International Union Against Cancer, Geneva), Vol. 48, pp. 4-23.

2. Edwards, A., Hammond, H. A., Lin, J., Caskey, C. T. & Chakraborty, R. (1992) *Genomics* **12**, 241-253.
3. Chamberlain, N. L., Driver, E. & Miesfeld, R. L. (1994) *Nucleic Acids Res.* **22**, 3181-3186.
4. Kazemi-Esfarjani, P., Trifiro, M. A. & Pinsky, L. (1995) *Hum. Mol. Genet.* **4**, 523-527.
5. LaSpada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E. & Fishbeck, K. H. (1991) *Nature (London)* **352**, 77-79.
6. Arbizu, T., Santamaria, J., Gomex, J. M., Quilez, A. & Serra, J. P. (1983) *J. Neurol. Sci.* **59**, 371-382.
7. Igarashi, S., Tanno, Y., Onodera, O., Yamazaki, M., Sato, S., *et al.* (1992) *Neurology* **42**, 2300-2302.
8. Irvine, R. A., Yu, M. C., Ross, R. K. & Coetzee, G. A. (1995) *Cancer Res.* **55**, 1937-1940.
9. Coetzee, G. A. & Ross, R. K. (1994) *J. Natl. Cancer Inst.* **86**, 872-873.
10. Woolf, C. M. (1960) *Cancer* **13**, 739-744.
11. Monroe, K. R., Yu, M., Kolonel, L. N., Coetzee, G. A., Wilkens, L. R., Ross, R. K. & Henderson, B. E. (1995) *Nat. Med.* **1**, 827-829.
12. Narod, S. A., Dupont, A., Cusan, L., Diamond, P., Gomez, J. L., Suburu, R. & Labrie, F. (1995) *Nat. Med.* **1**, 99-101.
13. Steinberg, G. D., Carter, B. S., Beaty, T. H., Childs, B. & Walsh, P. C. (1990) *Prostate* **17**, 337-347.
14. Hayes, R. B., Liff, J. M., Pottern, L. M., Greenberg, R. S., Schoenberg, J. B., *et al.* (1995) *Int. J. Cancer* **60**, 361-364.
15. Whittemore, A. S., Wu, A. H., Kolonel, L. N., John, E. M., Gallagher, R. P., Howe, G. R., West, D. W., Teh, C. Z. & Stamey, T. (1995) *Am. J. Epidemiol.* **141**, 732-740.
16. Steering Committee of the Physician's Health Study Research Group (1989) *N. Engl. J. Med.* **321**, 129-135.
17. Beahrs, O. H., Henson, D. E., Hutter, R. V. P. & Kennedy, B. J. (1992) *Manual for Staging of Cancer* (Lippincott, Philadelphia), 4th Ed.
18. Rothman, K. J. (1986) *Modern Epidemiology* (Little, Brown, Boston).
19. Yatani, R., Chigusa, I., Akazaki, K., Stemmermann, G. N., Welsh, R. A. & Correa, P. (1982) *Int. J. Cancer* **29**, 611-666.
20. Hovenian, M. S. & Deming, C. L. (1948) *Surg. Gynecol. Obstet.* **86**, 29-35.
21. Noble, R. L. (1977) *Cancer Res.* **37**, 1929-1933.
22. Pollard, M., Luckert, P. H. & Schmidt, M. A. (1982) *Prostate* **4**, 563-568.
23. Gann, P. H., Hennekens, C. H., Ma, J., Longcope, C. & Stampfer, M. J. (1996) *J. Natl. Cancer Inst.* **88**, 1118-1126.
24. Legro, R. S., Shahbahrani, B., Lobo, R. & Kovacs, B. W. (1994) *Obstet. Gynecol.* **83**, 701-706.
25. Ross, R. K. (1992) in *Accomplishments in Cancer Research, 1992*, General Motors Cancer Research Foundation, eds. Fortner, J. G. & Rhoads, J. E. (Lippincott, Philadelphia), pp. 219-228.
26. Hardy, D. O., Scher, H. I., Bogenreider, T., Sabbatini, P., Zhang, Z.-F., Nanus, D. M. & Catterall, J. F. (1996) *J. Endocrinol. Metab.* **81**, 4400-4405.
27. Brawn, P. N., Johnson, E. H., Kuhl, D. L., Riggs, M. W., Speights, V. O., Johnson, C. F., III, Pandya, P. P., Lind, M. L. & Bell, N. F. (1993) *Cancer* **71**, 2569-2573.
28. Moul, J. W., Sesterhenn, I. A., Connelly, R. R., Douglas, T., Srivastava, S., Mostofi, F. K. & McLeod, D. G. (1995) *J. Am. Med. Assoc.* **274**, 1277-1281.