How Strong Is the Association Between CAG and GGN Repeat Length Polymorphisms in the Androgen Receptor Gene and Prostate Cancer Risk?

Maurice P. Zeegers,[1,2,3] Lambertus A.L.M. Kiemeney,[4] Alan M. Nieder,[5] and Harry Ostrer[6]

[1]Department of Public Health and Epidemiology, School of Medicine, University of Birmingham, Birmingham, United Kingdom; [2]Department of General Practice, Comprehensive Cancer Institute Limburg, Catholic University of Leuven, Leuven, Belgium; [3]Department of Epidemiology, Faculty of Health Sciences, Maastricht University, Maastricht, Netherlands; [4]Department of Urology and Epidemiology, University Medical Centre Nijmegen, Nijmegen, Netherlands; [5]Department of Urology, School of Medicine, University of Miami, Miami, Florida and [6]Human Genetics Program, School of Medicine, New York University, New York, New York

Abstract

Objective: **Although narrative reviews have suggested an association between (CAG)*n* and (GGN)*n* polymorphisms in the *AR* gene and prostate cancer, it has never been quantified systematically. The purpose of this meta-analysis was to provide relative and absolute quantitative summary estimates with sufficient power.**
Method: **Publications were identified through database searches for epidemiologic studies published until February 2004. For each study, mean differences in repeat length between cases and controls were calculated as well as continuous odds ratios (OR) per one CAG or GGN repeat decrement and discrete ORs to compare prostate cancer risk in men with short CAG repeats (≤ 21 repeats) versus long CAG repeats (>21 repeats) and short GGN repeats (≤ 16 repeats) versus long GGN repeats (>16 repeats). The study-specific estimates were combined by random effects metaregression analyses.**
Results: **Nineteen case-control studies were included in this review comprising a total of 4,274 cases and 5,275 controls. Prostate cancer cases had on average 0.26 fewer CAG repeats and 0.09 fewer GGN repeats than controls. The continuous ORs of prostate cancer per one repeat decrement were 1.02 and 1.01 for CAG and GGN repeats, respectively. The summary discrete OR (95% confidence interval) were 1.19 (1.07-1.31) and 1.31 (1.06-1.61) for CAG and GGN repeat polymorphisms, respectively.**
Conclusion: **Although the presence of shorter repeats seemed to be modestly associated with prostate cancer risk, the absolute difference in number of repeats between cases and controls is <1 repeat. We question whether such a small difference is enough to yield measurable biological impact in prostate carcinogenesis.** (Cancer Epidemiol Biomarkers Prev 2004;13(11):1765–71)

Introduction

Prostate cancer is the most commonly diagnosed noncutaneous malignancy among men and the second leading cause of cancer-related deaths in men in Western countries (1). Until today, little is known about its etiology. Only age, family history, and ethnicity are considered to substantially influence prostate cancer risk.

Global ethnic variation contributes to the explanation of the substantial differences in prostate cancer incidence existing throughout the world, being highest in Western countries, especially among African American men, and lowest in developing countries. However, little is known about why certain populations are more susceptible for developing prostate cancer. Some have suggested that varying levels of androgens across ethnic groups may be responsible for these differences (2).

In our recent meta-analysis of 32 population-based studies, we have shown that family members of a prostate cancer patient experience a 2.46-fold increased risk of prostate cancer (3). The nature of this familial clustering is such that the risk is higher for family members with a brother with prostate cancer compared with those with an affected father.

Both observations of ethnic and familial clustering provided the rationale to study whether polymorphisms in the human androgen receptor (*AR*) gene located at the X chromosome (Xq11.2-q12) may explain this variation in susceptibility to prostate cancer. The *AR* gene comprises eight exons that encode for four functional domains including the amino-terminal transcription activation (transactivation) domain, the DNA binding domain, a hinge region, and the carboxyl-terminal ligand binding domain (4). The DNA binding domain and ligand binding domain are highly conserved through evolution, whereas the transactivation domain, which regulates expression of target genes, is highly polymorphic. Within this latter domain, there are three microsatellite trinucleotide repeats, two of which are polymorphic in length. Variability in the reported molecular size of the *AR* is due in part to the upstream

Received 4/20/04; revised 5/11/04; accepted 5/28/04.

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.

Requests for reprints: Maurice P. Zeegers, Department of Public Health and Epidemiology, School of Medicine, Public Health Building, University of Birmingham, Birmingham B15 2TT, United Kingdom. Phone: 44-121-414-6721; Fax: 44-121-414-7878. E-mail: M.P.Zeegers@bham.ac.uk

CAG and downstream GGN repeat polymorphisms, encoding polyglutamine and polyglycine regions, respectively (4-6).

Several studies have suggested that one or both of the trinucleotide repeat polymorphisms in *AR* may be associated with prostate cancer (7-41). However, to our knowledge, no meta-analysis of all previous studies on this relationship has been conducted thus far. Earlier narrative reviews (4, 42-47) have concluded that prostate cancer risk tend to be highest for men with shorter (CAG)*n* or (GGN)*n* alleles. Inconsistency between study results has been ascribed to small sample sizes (low power) or poor research strategies for those studies not observing an association. Although these reviews reported study-specific relative effects [i.e., relative risk or odds ratio (OR)], they did not attempt to quantify these in a summary measure. In addition, the reviews did not report on the absolute mean difference (MD) in number of CAG or GGN repeats between prostate cancer cases and controls.

The objective of this study is to increase power by meta-analyzing all epidemiologic studies up to February 2004 to provide both additive and relative quantitative summary estimates of the association between (CAG)*n* and (GGN)*n* repeat polymorphisms in the *AR* gene and prostate cancer risk and to evaluate changes in summary estimates by differences in study characteristics.

Materials and Methods

Search Strategy. Publications were independently identified by two authors (M.Z. and H.O.) through computerized Medline, Embase, Cancerlit, and Current Contents searches for retrospective or prospective (i.e., nested) studies published until February 2004 with no language restriction. The keywords used were a combination of prostat*, cancer, carcinoma, androgen receptor, polyglutamine, polyglycine, CAG, GGN, AR, and repeat. Furthermore, references cited in published original and review articles (4, 42-47) were examined until no further study was identified.

For inclusion in this analysis, the articles had to present data to estimate an OR or MD and the corresponding standard error (SE), which compares prostate cancer cases with unrelated controls on variation in CAG or GGN repeat length in the *AR* gene.

A distinction was made between an article and a study. In case the same study population was referred to in more than one article, these articles were considered as part of a single study. On the other hand, if an article presents data of different groups of research subjects, the results of these analyses were seen as separate studies.

Qualitative Data Extraction. Four reviewers (M.Z., L.K., A.N., and H.O.) extracted qualitative information from each article. They independently assessed the following qualitative items: design (retrospective case-control or nested case-control); population characteristics—ethnic label (Hispanic, Caucasian, Asian, or African) and geographic area (country of study); patient characteristics—number, age (years), and histologic grade prostate cancer (percentage advanced prostate cancer, T3-T4, M0; T0-T4, M1); and control characteristics—number, age (years), and disease status [healthy or benign prostatic hyperplasia (BPH)]. When continuous data were presented across different subgroups, we calculated a weighed average using the median of each subgroup weighed by the frequency of observations in these strata. For example, if a population's age distribution is presented in strata of 10 younger (61-65 years) men and 15 older men (66-70 years), the average age of the population is estimated to be $(10 \times 63 + 15 \times 68) / 25 = 66$ years. When disagreement existed on an item, it was discussed until consensus was reached. These qualitative items are used to provide covariables for inclusion in metaregression models to explore reasons for observed heterogeneity in results between the studies.

Quantitative Data Extraction. Two reviewers (M.Z. and H.O.) independently extracted data, allowing us to calculate mean number of repeats in cases and controls, MDs in numbers of repeats between cases and controls, ORs of prostate cancer for short versus long repeats, and ORs of prostate cancer per decrement of one repeat.

When mean number of repeats and corresponding SD for cases and controls were not reported, these were calculated using published tables with frequency distributions of repeats for cases and controls. When these tables were not published, these were reconstructed by measuring published frequency graphs. When no information was available to calculate the SD, it was estimated as a quarter of the repeat range or calculated by using the P of the unpooled t test comparison of means between cases and controls: $SD_{cases} = SD_{controls} = \frac{MD}{Z^* \sqrt{\frac{1}{n_{cases}} + \frac{1}{n_{controls}}}}$, where Z is the corresponding Z score of the P of the unpooled t test, n_{cases} is the number of cases, and $n_{controls}$ is the number of controls. The MD in both CAG and GGN repeats were calculated as $MD = m_{cases} - m_{controls}$, where m_{cases} is the mean repeat number among cases and $m_{controls}$ is the mean repeat number among controls. The SE of the MD was calculated by $SE_{MD} = \sqrt{\frac{VAR_{cases}}{n_{cases}} + \frac{VAR_{controls}}{n_{controls}}}$, where VAR_{cases} is the squared SD for cases and $VAR_{controls}$ is the squared SD for controls.

We extracted discrete ORs (OR_d) to compare prostate cancer risk in men with short CAG repeats (≤21 repeats) versus long CAG repeats (>21 repeats) and short GGN repeats (≤16 repeats) versus long GGN repeats (>16 repeats). The cutoff values were chosen because these were the pooled median value of repeats in the control populations. If these could not be recapitulated, other nearby cutoff points were used. Preferably, OR_d and corresponding SE of the $\ln(OR_d)$ were extracted directly from the original reports. For studies that reported separate ORs for more than two strata of numbers of repeats, we combined the repeat-specific ORs by using the prevalence of the controls as weight to calculate OR_d (48). If these were not available, two-way contingency tables were constructed for each study based on published or reconstructed exposure frequency distributions to calculate the OR_d. The SE of latter $\ln(OR_d)$ was calculated using the method of Woolf (49).

Continuous ORs (OR_c) per one CAG or GGN repeat decrement (OR_1) were preferably extracted directly from the published articles. ORs per one repeat increment were inverted. If OR_c and corresponding SE were not

presented, they were derived directly from the estimated $m_{controls}$, m_{cases}, and $VAR_{controls}$ using $OR_c = \exp\left(\frac{m_{controls}-m_{cases}}{VAR_{controls}}\right)$ and $SE_{ln(OR_c)} = \frac{SE_{MD}}{VAR_{controls}}$ according to the method of Chene and Thompson (50).

Statistical Analysis. To detect publication or related biases, we explored heterogeneity in funnel plots [i.e., plots of MD against their estimated precision (reciprocal of the variance)]. We examined funnel plot asymmetry visually and measured the degree of asymmetry by using the Egger et al. (51) unweighed regression asymmetry test.

Because of potential heterogeneity between studies, we estimated the summary MD (SMD), summary OR_d, and summary OR_c together with their corresponding 95% confidence intervals (CI) with random effects meta-analysis by using the Stata statistical software (52). The between-study variance was estimated by a noniterative procedure using a method of moments estimator. To explore reasons for heterogeneity in MD, we used three methods: (*a*) we calculated the I^2 statistics (i.e., the percentage of variability in point estimates that is due to heterogeneity; ref. 53) for each summary measure of CAG or GGN repeat polymorphisms and across specific strata based on different levels of the extracted study characteristics. Heterogeneity exists when I^2 within unstratified analyses is higher than within stratified analyses; we did (*b*) metaregression analyses by including the study characteristics as covariates in the regression model (statistical significance indicates an effect modifier) and (*c*) stratification.

Results

Study Characteristics. The search strategy revealed 35 articles reporting epidemiologic studies on the association between *AR* receptor polymorphisms and prostate cancer (7-41). Thirteen articles were excluded because prostate cancer cases and controls were genealogically related (n = 3; refs. 28-30), the association between *AR* receptor polymorphisms and prostate cancer characteristics was investigated among cases only (n = 8; refs. 31-37, 54), or no sufficient data could be extracted (n = 2; refs. 38, 39). Three articles were combined in the analysis because the same study was published more often (9-11). Three articles were considered to represent six separate studies because different study populations were investigated (20, 24, 40). The remaining 23 articles described 19 retrospective case-control studies (7, 9-12, 14-18, 20-26, 40) and 5 prospective case-control studies, which were nested in a cohort (8, 13, 19, 27, 41), comprising a total of 4,274 cases and 5,275 controls (Table 1).

All studies were published in English. The majority (79%) of studies were done in Caucasian populations (8-15, 17-27, 40, 41), one study was done among Hispanic men (7), two studies were done among African men (24, 40), and two studies used an Asian population (16, 20). Most studies were done in Western countries, including America (n = 12; refs. 7-11, 13, 15, 18, 25-27, 40, 41), Europe (n = 7; refs. 12, 17, 19, 20, 22, 23), and Australia (n = 1; ref. 21). Two studies were done in Asia (16, 20) and two studies were done in Africa (24). The average mean (SD) age at diagnosis of the patients across

Table 1. Study characteristics of published epidemiologic studies concerning the association between *AR* CAG and GGN repeat length polymorphism and prostate cancer

General information		Population		Cases		Controls	
Endnote ID	Design*	Ethnicity	Country	Age	Advanced†	Age	Disease
Balic et al. (7)	Retrospective	Hispanic	United States	64.0		57.0	Healthy
Chen et al. (8)	Prospective	Caucasian	United States	61.2	11.5	60.8	
Ingles et al. (9-11)	Retrospective	Caucasian	United States	57.8	46.0	58.2	Healthy
Gsur et al. (12)	Retrospective	Caucasian	Austria	65.9		66.5	BPH
Giovannucci et al. (13)	Prospective	Caucasian	United States		30.7		Healthy
Correa-Cerro et al. (14)	Retrospective	Caucasian	France/Germany	68.2		71.2	Healthy
Chang et al. (15)	Retrospective	Caucasian	United States	60.9		58.0	Healthy
Hsing et al. (16)	Retrospective	Asian	China	72.2	62.6	71.9	Healthy
Mononen et al. (17)	Retrospective	Caucasian	Finland	68.1	48.1		Healthy
Modugno et al. (18)	Retrospective	Caucasian	United States	68.9		73.6	Healthy
Latil et al. (19)	Prospective	Caucasian	France	70.5	69.8	71.7	Healthy
Ekman et al. (20)	Retrospective	Caucasian	Sweden	69.0		72+	BPH
	Retrospective	Asian	Japan	71.0		60+	BPH
Beilin et al. (21)	Retrospective	Caucasian	Australia	67.0	39.2	66.6	Healthy
Bratt et al. (22)	Retrospective	Caucasian	Sweden	70.2			Healthy
Edwards et al. (23)	Retrospective	Caucasian	United Kingdom	68.1	75.3		Healthy
Panz et al. (24)	Retrospective	Caucasian	Israel/South Africa	76.0	30.0		Healthy
	Retrospective	African	South Africa	68.0	30.0		Healthy
Stanford et al. (25)	Retrospective	Caucasian	United States	54.9	45.9	54.0	Healthy
Hakimi et al. (26)	Retrospective	Caucasian	United States	62.1	42.4		Healthy
Platz et al. (27)	Prospective	Caucasian	United States	62.0	46.6		Healthy
Dos Santos et al. (40)	Retrospective	Caucasian	Brazil	65.0		58.0	Healthy
	Retrospective	African	Brazil	65.0		58.0	Healthy
Visvanathan et al. (41)	Prospective	Caucasian	United States	66.1		66.0	Healthy

*Retrospective or prospective (=nested) case-control study.
†Percentage advanced prostate cancer (T3-T4, M0; T0-T4, M1).

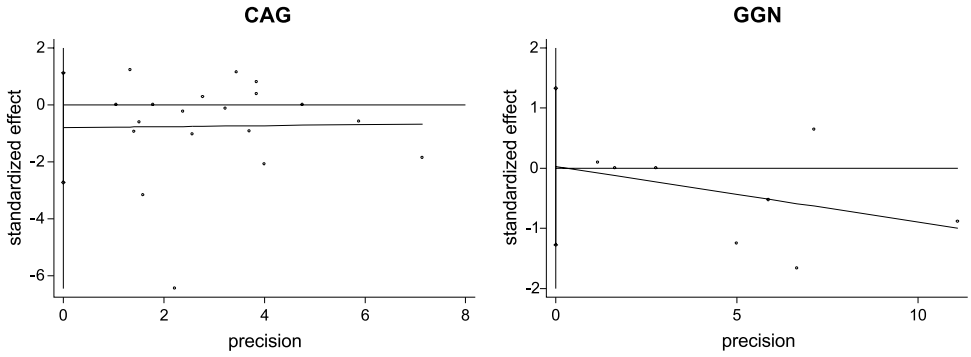

Figure 1. Publication bias plot for CAG and GGN repeat polymorphisms. Standardized effect estimates versus precision along with the regression line and the 95% CI about the intercept. Failure of this 95% CI to include zero indicates asymmetry in the funnel plot and may give evidence of publication bias. Guidelines include $x = 0$ and $y = 0$.

all studies was 66.2 (4.8) years ranging from 54.9 to 76.0 years. On average, 42.4% of the cases suffered from advanced prostate cancer. The average percentage of cases with a Gleason score ≥7 across all studies was 32.6% (data not shown). The average mean (SD) age of controls was 63.7 (6.7) years ranging from 54.0 to 73.6 years. Most studies selected healthy controls, although three studies selected controls with BPH (12, 20; Table 1).

Publication Bias. We could not identify heterogeneity in funnel plots neither visually (Fig. 1) nor in terms of statistical significance (P = 0.39 and 0.96 for CAG and GGN repeats, respectively).

CAG Repeats. Nineteen studies conveyed data on the mean number and SD of CAG repeats among cases and controls (7, 8, 12-17, 19-27, 40; Fig. 2). Statistical pooling showed that prostate cancer cases seemed to have on average 0.26 fewer CAG repeats than controls (SMD, −0.26; 95% CI, −0.53 to 0.00). The summary OR_d (95% CI) comparing men with ≤21 CAG repeats to men with >21 CAG repeats was 1.19 (1.07-1.31) based on the results of 22 studies (7-26, 40, 41). The OR_c (95% CI) of prostate cancer per one CAG decrement was 1.02 (0.99-1.06) based on the results from the same studies that were used to calculate the SMD (7, 8, 12-17, 19-27, 40; Fig. 2).

GGN Repeats. The calculation of the SMD for the number of GGN repeats was based on eight studies (8, 14-16, 23, 25-27; Fig. 2). On average, prostate cancer cases had 0.09 fewer GGN repeats than controls (SMD, −0.09; 95% CI, −0.20 to 0.03). Men with ≤16 GGN repeats seemed to have a 31% higher risk of prostate cancer than

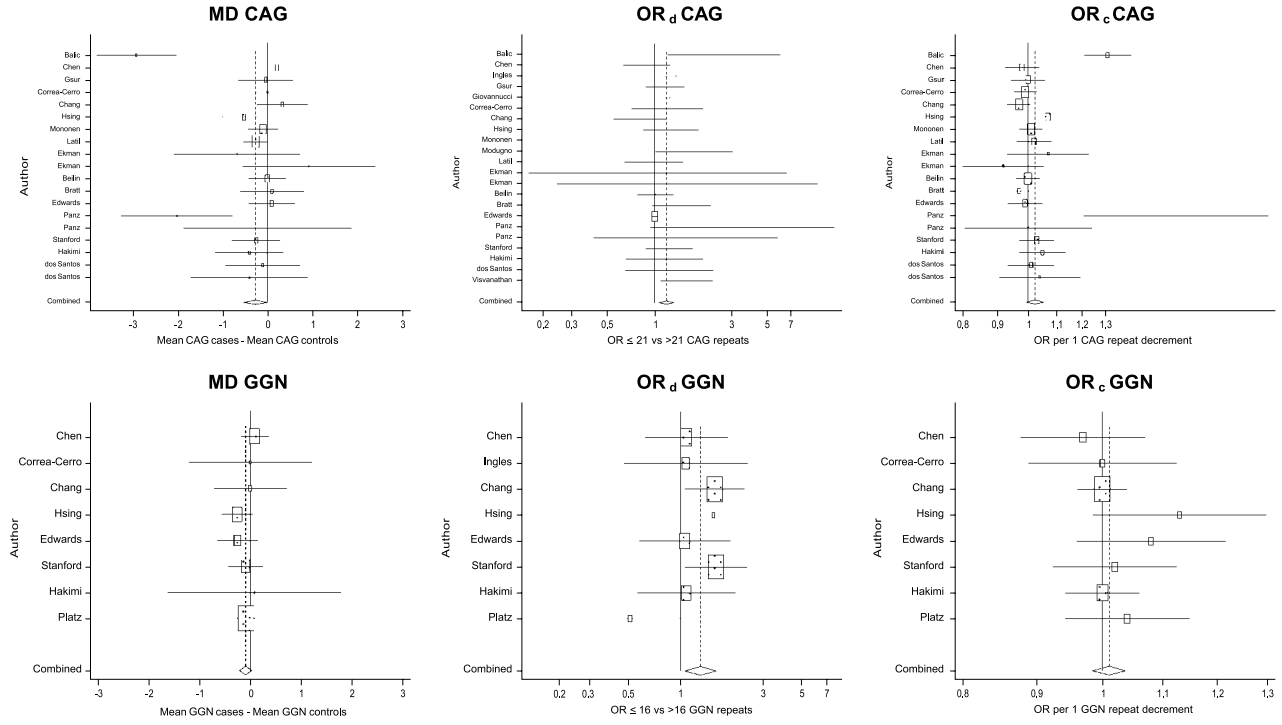

Figure 2. Study-specific estimates (*boxes*) and corresponding 95% CI (*horizontal lines*) for the influence of CAG and GGN repeat polymorphisms on prostate cancer risk. Box sizes are proportional to study weight. Summary statistics and 95% CI are represented by *diamonds*. Guidelines include the null and summarized effects.

Table 2. I^2 for CAG repeats both unstratified and stratified for specific study characteristics, *P* of interaction for individual study characteristics resulting from metaregresson analysis, and MD and corresponding 95% CI per stratum

	No. studies	MD (95% CI)	I^2 (%)	$P_{interaction}$
Unstratified	19	−0.26 (−0.53 to 0.00)	69	
Stratified				
Design				
Prospective	2	−0.08 (−0.53 to 0.37)	61	0.52
Retrospective	17	−0.31 (−0.63 to 0.01)	71	
Ethnicity of study population				
Caucasian	14	−0.10 (−0.26 to 0.07)	22	0.02
Other	5	−0.67 (−1.94 to 0.60)	86	
Country of study				
United States	7	−0.46 (−1.17 to 0.24)	86	0.31
Other	12	−0.16 (−1.37 to 0.05)	32	
Mean age of cases (y)				
<65	5	−0.55 (−1.46 to 0.35)	90	0.37
≥65	14	−0.17 (−0.35 to 0.02)	20	
Advanced prostate cancer				
<50%	7	−0.18 (−0.50 to 0.13)	48	0.90
≥50%	12	−0.29 (−0.69 to 0.12)	80	
Mean age of controls (y)				
<65	6	−0.60 (−1.71 to 0.51)	93	0.41
≥65	13	−0.18 (−0.37 to 0.02)	27	
Disease status of controls*				
Healthy	15	−0.35 (−0.65 to −0.04)	74	0.41
BPH	3	−0.01 (−0.63 to 0.64)	17	

*The disease status of the controls in one study (8) could not be recapitulated.

men with >16 GGN repeats (OR_d, 1.31; 95% CI, 1.06-1.61). This estimate is based on the results of eight studies (8-11, 15, 16, 23, 25-27). The OR_c (95% CI) of prostate cancer per one GGN decrement was 1.01 (0.98-1.04) based on the results from the same studies that were used to calculate the SMD (8, 14-16, 23, 25-27; Fig. 2).

Heterogeneity. We further examined heterogeneity in study results by categories of study characteristics including study design, ethnicity of study population, country of study, age of cases and controls, severity of disease, and type of disease status of controls. The definitions of the specific categories can be found in Table 2. The analyses were limited to studies reporting the association between number of CAG repeats and prostate cancer risk only, because the data on GGN repeats were too sparse. Except for the ethnicity of the study population ($P = 0.02$), none of the study characteristics seemed to be a statistically significant interaction term in metaregression analyses. However, quantitative differences in stratified MDs in CAG repeats were found. The MD in number of CAG repeats between cases and controls seemed to be larger in retrospective (−0.31) versus prospective (−0.08) case-control studies, non-Caucasian (−0.67) versus Caucasian (−0.10) study populations, the United States (−0.046) versus other countries (−0.16), younger (−0.55) versus older (−0.17) cases, more (−0.29) versus less (−0.18) cases with advanced prostate cancer, younger (−0.48) versus older (−0.18) controls, and healthy (−0.35) versus BPH (−0.01) controls. However, within all subgroups, still a considerable amount of the variation (17-93%) could be explained by random heterogeneity. Furthermore, none of the stratified results showed a meaningful difference in CAG repeat between cases and controls. All differences were <1 repeat (Table 2).

Discussion

Several biological observations have led to the expectation that *AR* repeat polymorphisms may be associated with susceptibility to developing prostate cancer. First, it is known that androgens play a role in human prostate cancer growth, because dogs and males castrated before puberty do not get prostate cancer and because androgen-ablative therapy initially inhibits prostate tumor growth (46, 55). Second, it has been observed that the length of the CAG repeat region is inversely related to the transactivation activity on the receptor and the binding affinity for androgens (56-58). Irvine et al. (59) reported that increased CAG length inhibited both basal and coactivator-mediated *AR* transactivation activity in cultured prostate epithelial cells that were cotransfected with *AR* constructs of up to 42 repeats and p160 coactivator expression vectors. Therefore, decreased repeat length may make the prostate more vulnerable to chronic androgen stimulation leading to increased proliferative activity, which in turn may increase the rate of somatic mutations (46). Third, *AR* repeat polymorphisms have been shown to be associated with the incidence of other androgen-related clinical conditions; specifically, a high number of repeats seem to adversely influence fertility and spermatogenesis (60) and bone density (61), and fewer repeats are associated with increased risk of baldness (62) and benign prostatic hyperplasia (63-66), although this has not always been confirmed (67).

In this meta-analysis, we conclude that the relative epidemiologic associations derived from previous literature confirm these biological expectations. We have shown that men with shorter CAG and GGN repeats experience a 1.19- and 1.31-fold increased risk of prostate

cancer, respectively, although these increased risks should be interpreted as being very small. It should be noted that it is possible that CAG and GGC repeats are in linkage disequilibrium and it should therefore be considered that effects seen at one microsatellite might reflect effects at the other.

The results from the heterogeneity analyses were to some extent consistent with previous literature. For example, the association between CAG repeat polymorphism and prostate cancer seemed to be somewhat larger when controls were compared with cases with more advanced prostate cancer or with cancer that was diagnosed at an earlier age. According to Giovannucci (44), this could suggest that a heightened state of androgenicity directly influences a pool of relatively aggressive, early-onset androgen-driven prostate cancers. Possibly, prostate cancers that occur at older ages may be less driven by androgenicity and more related to pathologic processes such as oxidative insults (44). In addition, the MD in number of CAG repeats between cases and controls seemed to be larger for non-Caucasians and study populations from developing countries. Black males have a population average of 18 repeats, whereas Caucasian men have on average 20 repeats (47). Combining this with the fact that there is an inverse relationship between CAG length and *AR* transactivation function, such that the shorter the CAG, the better the *AR* is at activating gene transcription (68), Coetzee and Ross (57) hypothesized that the racial differences seen in prostate cancer susceptibility may be due in part to variability in the *AR* CAG repeat length (4). However, because the association between *AR* repeat length polymorphisms and prostate cancer is small, other factors must be involved as well.

The incidence of prostate cancer has increased in several countries because the availability of prostate-specific antigen testing in the late 1980s in the United States and in the early 1990s in other countries. As pointed out by Giovannucci (44), pre–prostate-specific antigen era patients might differ significantly from post–prostate-specific antigen era patients. However, in this meta-analysis, all studies have been conducted after the introduction of prostate-specific antigen screening. Because prostate-specific antigen screening of the general population is of more common use in the United States and less often recommended and carried out by other countries, the larger SMD in CAG repeat polymorphism between prostate cancer cases in the United States compared with other countries may be attributed to differences in screening policy.

Although previous epidemiologic studies have consistently reported that the magnitude of relative association between *AR* repeat polymorphisms and prostate cancer is relatively limited, most narrative reviews to date have suggested a positive association between shorter CAG or GGN repeat lengths and prostate cancer. Some have argued that even a modest increased risk could still be of importance from a public health perspective because of the large population frequency of short repeat lengths (46). Based on this reasoning, Nelson and Witte (46) suggested that this potentially low-penetrance, high-frequency polymorphism could theoretically account for many more prostate cancer cases than a high-penetrance, low-frequency mutation for which the etiologic fraction is estimated to be 9%. This may be true on a population scale for the GGN polymorphism but not for the CAG polymorphism. We have estimated that the etiologic fraction is 13.4% and 8.7% for GGN and CAG repeat polymorphisms, respectively, assuming a short repeat frequency of 50% (because median cutoff values were used) and the reported summary OR_d.

However, our analyses also show that although cases seemed to have shorter repeats and shorter repeats seemed to be modestly associated with prostate cancer risk on a relative scale, the absolute difference in number of repeats between cases and controls is <1 repeat. The SMDs in number of CAG and GGN repeats were −0.27 and −0.09, respectively. Although the direction of results in the heterogeneity analysis seemed to be consistent with previous literature, none of the stratified differences between cases and controls exceeded one repeat. Based on these results, we question whether an average MD of <1 repeat is enough to yield measurable biological impact in prostate carcinogenesis.

References

1. Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics, 1999. CA Cancer J Clin 1999;49:8–31, 1.
2. Latil A, Lidereau R. Genetic aspects of prostate cancer. Virchows Arch 1998;432:389–406.
3. Zeegers MP, Jellema A, Ostrer H. Empiric risk of prostate carcinoma for relatives of patients with prostate carcinoma: a meta-analysis. Cancer 2003;97:1894–903.
4. Montgomery JS, Price DK, Figg WD. The androgen receptor gene and its influence on the development and progression of prostate cancer. J Pathol 2001;195:138–46.
5. Choong CS, Wilson EM. Trinucleotide repeats in the human androgen receptor: a molecular basis for disease. J Mol Endocrinol 1998;21:235–57.
6. Bentel JM, Tilley WD. Androgen receptors in prostate cancer. J Endocrinol 1996;151:1–11.
7. Balic I, Graham ST, Troyer DA, et al. Androgen receptor length polymorphism associated with prostate cancer risk in Hispanic men. J Urol 2002;168:2245–8.
8. Chen C, Lamharzi N, Weiss NS, et al. Androgen receptor polymorphisms and the incidence of prostate cancer. Cancer Epidemiol Biomarkers Prev 2002;11:1033–40.
9. Ingles SA, Ross RK, Yu MC, et al. Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. J Natl Cancer Inst 1997;89:166–70.
10. Xue W, Irvine RA, Yu MC, Ross RK, Coetzee GA, Ingles SA. Susceptibility to prostate cancer: interaction between genotypes at the androgen receptor and prostate-specific antigen loci. Cancer Res 2000;60:839–41.
11. Irvine RA, Yu MC, Ross RK, Coetzee GA. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. Cancer Res 1995;55:1937–40.
12. Gsur A, Preyer M, Haidinger G, et al. Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. Carcinogenesis 2002;23:1647–51.
13. Giovannucci E, Stampfer MJ, Krithivas K, et al. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. Proc Natl Acad Sci U S A 1997;94:3320–3.
14. Correa-Cerro L, Wohr G, Haussler J, et al. (CAG)*n*CAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French-German population. Eur J Hum Genet 1999;7:357–62.
15. Chang BL, Zheng SL, Hawkins GA, et al. Polymorphic GGC repeats in the androgen receptor gene are associated with hereditary and sporadic prostate cancer risk. Hum Genet 2002;110:122–9.
16. Hsing AW, Gao YT, Wu G, et al. Polymorphic CAG and GGN repeat lengths in the androgen receptor gene and prostate cancer risk: a population-based case-control study in China. Cancer Res 2000;60:5111–6.
17. Mononen N, Ikonen T, Autio V, et al. Androgen receptor CAG polymorphism and prostate cancer risk. Hum Genet 2002;111:166–71.

18. Modugno F, Weissfeld JL, Trump DL, et al. Allelic variants of aromatase and the androgen and estrogen receptors: toward a multigenic model of prostate cancer risk. Clin Cancer Res 2001;7: 3092–6.
19. Latil AG, Azzouzi R, Cancel GS, et al. Prostate carcinoma risk and allelic variants of genes involved in androgen biosynthesis and metabolism pathways. Cancer 2001;92:1130–7.
20. Ekman P, Gronberg H, Matsuyama H, Kivineva M, Bergerheim US, Li C. Links between genetic and environmental factors and prostate cancer risk. Prostate 1999;39:262–8.
21. Beilin J, Harewood L, Frydenberg M, et al. A case-control study of the androgen receptor gene CAG repeat polymorphism in Australian prostate carcinoma subjects. Cancer 2001;92:941–9.
22. Bratt O, Borg A, Kristoffersson U, Lundgren R, Zhang QX, Olsson H. CAG repeat length in the androgen receptor gene is related to age at diagnosis of prostate cancer and response to endocrine therapy, but not to prostate cancer risk. Br J Cancer 1999;81:672–6.
23. Edwards SM, Badzioch MD, Minter R, et al. Androgen receptor polymorphisms: association with prostate cancer risk, relapse and overall survival. Int J Cancer 1999;84:458–65.
24. Panz VR, Joffe BI, Spitz I, Lindenberg T, Farkas A, Haffejee M. Tandem CAG repeats of the androgen receptor gene and prostate cancer risk in black and white men. Endocrine 2001;15:213–6.
25. Stanford JL, Just JJ, Gibbs M, et al. Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. Cancer Res 1997;57:1194–8.
26. Hakimi JM, Schoenberg MP, Rondinelli RH, Piantadosi S, Barrack ER. Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. Clin Cancer Res 1997;3:1599–608.
27. Platz EA, Giovannucci E, Dahl DM, et al. The androgen receptor gene GGN microsatellite and prostate cancer risk. Cancer Epidemiol Biomarkers Prev 1998;7:379–84.
28. Ho GY, Knapp M, Freije D, et al. Transmission/disequilibrium tests of androgen receptor and glutathione *S*-transferase π variants in prostate cancer families. Int J Cancer 2002;98:938–42.
29. Lange EM, Chen H, Brierley K, et al. The polymorphic exon 1 androgen receptor CAG repeat in men with a potential inherited predisposition to prostate cancer. Cancer Epidemiol Biomarkers Prev 2000;9:439–42.
30. Miller EA, Stanford JL, Hsu L, Noonan E, Ostrander EA. Polymorphic repeats in the androgen receptor gene in high-risk sibships. Prostate 2001;48:200–5.
31. Bennett CL, Price DK, Kim S, et al. Racial variation in CAG repeat lengths within the androgen receptor gene among prostate cancer patients of lower socioeconomic status. J Clin Oncol 2002;20:3599–604.
32. Cude KJ, Montgomery JS, Price DK, et al. The role of an androgen receptor polymorphism in the clinical outcome of patients with metastatic prostate cancer. Urol Int 2002;68:16–23.
33. Hardy DO, Scher HI, Bogenreider T, et al. Androgen receptor CAG repeat lengths in prostate cancer: correlation with age of onset. J Clin Endocrinol Metab 1996;81:4400–5.
34. Mir K, Edwards J, Paterson PJ, Hehir M, Underwood MA, Bartlett JM. The CAG trinucleotide repeat length in the androgen receptor does not predict the early onset of prostate cancer. BJU Int 2002;90:573–8.
35. Nam RK, Elhaji Y, Krahn MD, et al. Significance of the CAG repeat polymorphism of the androgen receptor gene in prostate cancer progression. J Urol 2000;164:567–72.
36. Shibata A, Garcia MI, Cheng I, et al. Polymorphisms in the androgen receptor and type II 5α-reductase genes and prostate cancer prognosis. Prostate 2002;52:269–78.
37. Watanabe M, Ushijima T, Shiraishi T, et al. Genetic alterations of androgen receptor gene in Japanese human prostate cancer. Jpn J Clin Oncol 1997;27:389–93.
38. Platz EA, Rimm EB, Willett WC, Kantoff PW, Giovannucci E. Racial variation in prostate cancer incidence and in hormonal system markers among male health professionals. J Natl Cancer Inst 2000; 92:2009–17.
39. Suzuki H, Akakura K, Komiya A, et al. CAG polymorphic repeat lengths in androgen receptor gene among Japanese prostate cancer patients: potential predictor of prognosis after endocrine therapy. Prostate 2002;51:219–24.
40. Santos ML, Sarkis AS, Nishimoto IN, Nagai MA. Androgen receptor CAG repeat polymorphism in prostate cancer from a Brazilian population. Cancer Detect Prev 2003;27:321–6.
41. Visvanathan K, Helzlsouer KJ, Boorman DW, et al. Association among an ornithine decarboxylase polymorphism, androgen receptor gene (CAG) repeat length and prostate cancer risk. J Urol 2004; 171:652–5.
42. Buchanan G, Irvine RA, Coetzee GA, Tilley WD. Contribution of the androgen receptor to prostate cancer predisposition and progression. Cancer Metastasis Rev 2001;20:207–23.
43. Coetzee G, Irvine R. Size of the androgen receptor CAG repeat and prostate cancer: does it matter? J Clin Oncol 2002;20:3572–3.
44. Giovannucci E. Is the androgen receptor CAG repeat length significant for prostate cancer? Cancer Epidemiol Biomarkers Prev 2002;11:985–6.
45. Kantoff P, Giovannucci E, Brown M. The androgen receptor CAG repeat polymorphism and its relationship to prostate cancer. Biochim Biophys Acta 1998;1378:C1–5.
46. Nelson KA, Witte JS. Androgen receptor CAG repeats and prostate cancer. Am J Epidemiol 2002;155:883–90.
47. Price DK, Franks ME, Figg WD. Genetic variations in the vitamin D receptor, androgen receptor and enzymes that regulate androgen metabolism. J Urol 2004;171:S45–9; discussion S49.
48. Tan FE, Zeegers MP. An asymptotically unbiased estimator of exposed versus non-exposed odds ratio from reported dose-response data. Stat Methods Med Res 2001;10:311–23.
49. Woolf B. On estimation the relation between blood group and disease. Ann Hum Genet 1954;19:251–3.
50. Chene G, Thompson SG. Methods for summarizing the risk associations of quantitative variables in epidemiologic studies in a consistent form. Am J Epidemiol 1996;144:610–21.
51. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.
52. Stata statistical software: release 6.0. College Station (TX): Stata Corporation.
53. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58.
54. Koivisto PA, Hyytinen ER, Matikainen M, Tammela TL, Ikonen T, Schleutker J. Germline mutation analysis of the androgen receptor gene in Finnish patients with prostate cancer. J Urol 2004;171:431–3.
55. Wilding G. The importance of steroid hormones in prostate cancer. Cancer Surv 1992;14:113–30.
56. Feldman D. Androgen and vitamin D receptor gene polymorphisms: the long and short of prostate cancer risk. J Natl Cancer Inst 1997; 89:109–11.
57. Coetzee GA, Ross RK. Re: Prostate cancer and the androgen receptor. J Natl Cancer Inst 1994;86:872–3.
58. Mhatre AN, Trifiro MA, Kaufman M, et al. Reduced transcriptional regulatory competence of the androgen receptor in X-linked spinal and bulbar muscular atrophy. Nat Genet 1993;5:184–8.
59. Irvine RA, Ma H, Yu MC, Ross RK, Stallcup MR, Coetzee GA. Inhibition of p160-mediated coactivation with increasing androgen receptor polyglutamine length. Hum Mol Genet 2000;9:267–74.
60. Yoshida KI, Yano M, Chiba K, Honda M, Kitahara S. CAG repeat length in the androgen receptor gene is enhanced in patients with idiopathic azoospermia. Urology 1999;54:1078–81.
61. Zitzmann M, Brune M, Kornmann B, Gromoll J, Junker R, Nieschlag E. The CAG repeat polymorphism in the androgen receptor gene affects bone density and bone metabolism in healthy males. Clin Endocrinol (Oxf) 2001;55:649–57.
62. Sawaya ME, Shalita AR. Androgen receptor polymorphisms (CAG repeat lengths) in androgenetic alopecia, hirsutism, and acne. J Cutan Med Surg 1998;3:9–15.
63. Giovannucci E, Stampfer MJ, Chan A, et al. CAG repeat within the androgen receptor gene and incidence of surgery for benign prostatic hyperplasia in U.S. physicians. Prostate 1999;39:130–4.
64. Giovannucci E, Platz EA, Stampfer MJ, et al. The CAG repeat within the androgen receptor gene and benign prostatic hyperplasia. Urology 1999;53:121–5.
65. Mitsumori K, Terai A, Oka H, et al. Androgen receptor CAG repeat length polymorphism in benign prostatic hyperplasia (BPH): correlation with adenoma growth. Prostate 1999;41:253–7.
66. Shibata A, Stamey TA, McNeal JE, Cheng I, Peehl DM. Genetic polymorphisms in the androgen receptor and type II 5α-reductase genes in prostate enlargement. J Urol 2001;166:1560–4.
67. Bousema JT, Bussemakers MJ, van Houwelingen KP, et al. Polymorphisms in the vitamin D receptor gene and the androgen receptor gene and the risk of benign prostatic hyperplasia. Eur Urol 2000;37:234–8.
68. Chamberlain NL, Driver ED, Miesfeld RL. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. Nucleic Acids Res 1994;22: 3181–6.