Int. J. Cancer: **87,** 637–643 (2000)

Publication of the International Union Against Cancer

ANDROGEN RECEPTOR EXON 1 CAG REPEAT LENGTH AND RISK OF OVARIAN CANCER

Amanda B. SPURDLE[1]*, Penelope M. WEBB[2], Xiaoqing CHEN[1], Nicholas G. MARTIN[3], Graham G. GILES[4], John L. HOPPER[5] and Georgia CHENEVIX-TRENCH[1]

[1]*Cancer Unit, Joint Experimental Oncology Programme, The Queensland Institute of Medical Research and The University of Queensland, Brisbane, Queensland, Australia*
[2]*Department of Social and Preventive Medicine, University of Queensland Medical School, Herston, Queensland, Australia*
[3]*Epidemiology Unit, The Queensland Institute of Medical Research and The University of Queensland, Brisbane, Queensland, Australia*
[4]*Cancer Epidemiology Centre, Anti-Cancer Council of Victoria, Australia*
[5]*Centre for Genetic Epidemiology, The University of Melbourne, Carlton, Australia*

Epidemiological studies indicate that ovarian cancer is an endocrine-related tumour. We conducted a case-control comparison to assess the androgen receptor (AR) exon 1 polymorphic CAG repeat length (CAG_n) as a risk factor for epithelial ovarian cancer. AR CAG_n was determined for 319 case subjects with ovarian adenocarcinoma and 853 unaffected control subjects (comprising 300 unrelated adult female monozygotic twins, and 553 adult females sampled randomly from the population using the electoral rolls). The CAG_n distributions of case subjects and control subjects were compared as a continuum, and by dichotomising alleles according to different CAG_n cut-points. Logistic regression was used to calculate age-adjusted odds ratio (OR) estimates. Analyzed as a continuous variable, there was no difference between case subjects and control subjects for the smaller, larger or average allele sizes of the CAG_n genotype, before or after adjusting for age. The mean (95% CI) for the average CAG_n was 22.0 (21.8–22.2) for case subjects and 22.0 (21.9–22.1) for control subjects ($p > .9$). Analysis of CAG_n as a dichotomous variable showed no difference between case subjects and control subjects for the median cutpoint (≥ 22), or for another cut-point previously reported to act as a modifier of breast cancer risk (≥ 29). Our data provide no evidence for an association between ovarian cancer risk and the genotype defined by the AR exon 1 CAG_n polymorphism, although we cannot exclude small effects, or threshold effects in a small subgroup. *Int. J. Cancer* 87:637–643, 2000.
© *2000 Wiley-Liss, Inc.*

Ovarian cancer is the main cause of death among women with gynaecological malignancies, and the lifetime risk in Australian women is 1 in 99 (AIHW and AACR, 1998). Other than age, family history is the strongest risk factor for ovarian cancer (Parazzini *et al.*, 1991; Purdie *et al.*, 1995). For example, in Australia, having one first-degree relative with ovarian cancer was found to be associated with a fourfold increase in risk of ovarian cancer (Purdie *et al.*, 1995). Hereditary ovarian cancer can be caused by mutations in the breast cancer susceptibility genes BRCA1 and BRCA2, or in the mismatch repair genes hMSH2 and hMLH1 (reviewed by Boyd and Rubin, 1997). However, mutations in these genes are unlikely to contribute greatly to the aetiology of ovarian cancer in general, since most ovarian cancer cases are "sporadic" in that they do not appear to have a family history of the disease, and likewise the vast majority (~99%) of Australian women with ovarian cancer would not be classified as "high-risk" familial cases. Furthermore, the rarity of mutations in BRCA1, BRCA2 and the mismatch repair genes would suggest that they are unlikely to explain more than a small proportion of familial aggregation in ovarian cancer, and BRCA1 mutations have been shown to account for only about 5% of ovarian cancer cases diagnosed before the age of 70 years in a population-based UK study (Stratton *et al.*, 1997), while a recent UK study of familial ovarian cancer detected BRCA1 and BRCA2 mutations in only 20% of families with two cases of ovarian cancer (Gayther *et al.*, 1999). It is thus likely that common "low-risk" allelic variants in these or other genes account for at least some predisposition to ovarian cancers occurring in the general population. In an attempt to identify such low-risk ovarian cancer susceptibility genes, we are using the candidate gene approach to compare large samples of women with ovarian cancer and control subjects.

Epidemiological studies indicate that ovarian cancer is an endocrine-related tumour (Parazzini *et al.*, 1991), and androgens have been implicated in the aetiology of the disease (Risch, 1998). Elevated levels of androstenedione and dehydroepiandrosterone have been observed in case subjects (Helzlsouer *et al.*, 1995), androgen receptors have been identified within epithelial cells of normal ovaries (Al-Timimi *et al.*, 1985), and animal models indicate that testosterone stimulates the growth of ovarian surface papillomas and cystadenomas *in vivo* (Silva *et al.*, 1997).

The androgen receptor (AR) gene is involved in various pathways, including the differentiation, development and regulation of cell growth. A role in cancer predisposition is suggested by reported associations between prostate cancer risk and the length of the polymorphic exon 1 CAG repeat (CAG_n) within the AR transactivation domain (Irvine *et al.*, 1995; Giovanucci *et al.*, 1997; Hakimi *et al.*, 1997; Ingles *et al.*, 1997; Stanford *et al.*, 1997). From the two largest research studies, one with 269 high-grade prostate cancer case subjects and 588 control subjects found a relative risk of 2.1 (95% confidence interval [CI] = 1.1–4.0) for $CAG_n < 19$ (Giovanucci *et al.*, 1997), while a second study of 281 prostate cancer case subjects and 246 control subjects found a relative risk of 2.2 (95% CI = 1.1–4.7) for $CAG_n < 22$ in a subgroup of relatively thin individuals (body mass index < 24.4) (Stanford *et al.*, 1997). Furthermore, early onset prostate cancer case subjects have been reported to have shorter repeat lengths (Hardy *et al.*, 1996), and $CAG_n < 22$ also appears to be associated with an increased risk of benign prostatic hyperplasia (Giovanucci *et al.*, 1999).

Biological significance of the CAG repeat length variation is suggested by *in vitro* studies, which have demonstrated that smaller repeat lengths exhibit greater transactivation capabilities (Chamberlain *et al.*, 1994). *In vivo*, greatly expanded CAG_n (> 39) has been shown to be associated with spinal and bulbar muscular atrophy (SBMA) (La Spada *et al.*, 1991). The biological importance is emphasized at the extreme lengths within the SBMA range, with an increase in repeat length correlating with younger age at onset of this disorder, and also with the likelihood of clinical

Grant sponsor: National Health and Medical Research Council; Grant number: 98 1311.

*Correspondence to: Amanda B. Spurdle, Ph.D., Cancer Unit, Queensland Institute of Medical Research, P.O. Royal Brisbane Hospital, Queensland, 4029, Australia. Fax: +617 3362 0105.
E-mail: mandyS@qimr.edu.au

Received 21 December 1999; Accepted 20 March 2000

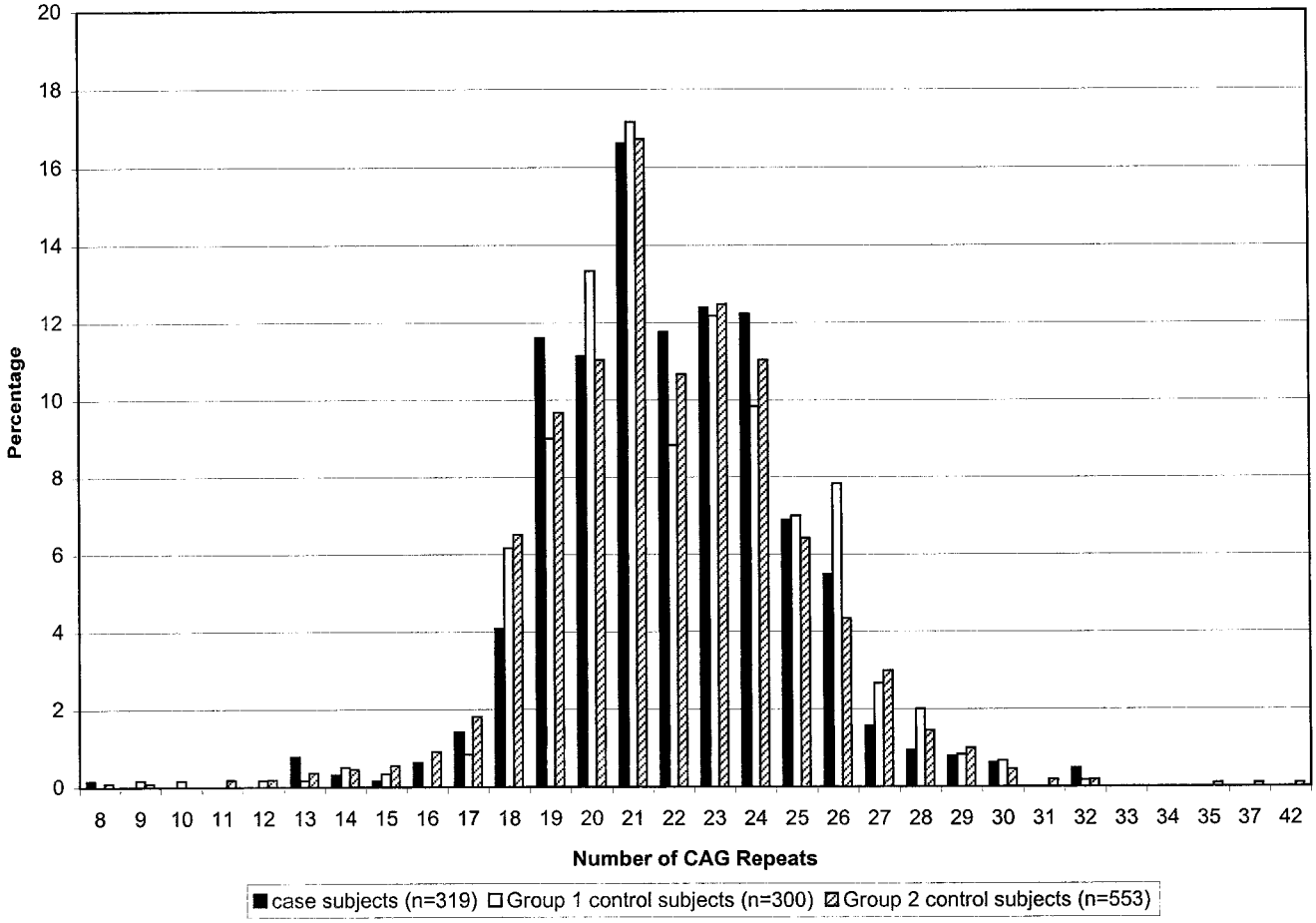

FIGURE 1 – Androgen receptor exon 1 CAG_n allele frequency distribution in ovarian cancer case subjects and control subjects.

manifestation of gynaecomastia (MacLean *et al.*, 1995). Within the normal range of repeat length variation, the biological importance is indicated not only by the prostate cancer studies detailed above, but also by a more recent report of an association between $CAG_n > 28$ and an increased risk of impaired spermatogenesis (Tut *et al.*, 1997), and by a retrospective study of 304 BRCA1 mutation carriers which suggested an association between $CAG_n \geqslant 28$ and earlier age at onset of BRCA1-associated breast cancer (Rebbeck *et al.*, 1999). However, we found that CAG_n had no effect on breast cancer risk in Australian women before the age of 40 years (Spurdle *et al.*, 1999).

We have undertaken a large case-control comparison to assess the AR exon 1 CAG repeat length polymorphism as a risk factor for epithelial ovarian cancer in Australian women. To the best of our knowledge, there have been no previous reports on this polymorphism and ovarian cancer.

MATERIAL AND METHODS

Subjects

Unselected women with ovarian adenocarcinoma (case subjects) were ascertained as incident cases from the Royal Brisbane Hospital, Queensland, Australia, during the period 1985–1996, and the series of 319 case subjects represented an inclusion rate of approximately 52%. Clinical information available for case-subjects included cancer form (benign, low malignant potential [LMP], or invasive), tumour histology, stage and grade. The series comprised 34 benign, 43 LMP and 241 invasive tumours, and a single tumour of unknown form. There were 196 serous, 38 mucinous, 32 endometrioid, 16 clear cell carcinoma, 8 mixed mullerian, 7 mixed, 4 undifferentiated and one Brenner tumour(s), as well as 17 of unknown histology. Patients were staged at laparotomy in accordance with the recommendations of the International Federation of Gynaecology and Obstetrics (FIGO) (Pettersson, 1988). Of the 234 invasive tumours of known stage, there were 27, 21, 159 and 27 at FIGO stages 1, 2, 3 and 4, respectively. The 219 invasive tumours for which grade was known included 82, 34 and 103 tumours classified as grade 1–2, 2/3 and 3–4 respectively. The age at diagnosis of case subjects ranged from 21–95 years, with an average of 59.8 years (standard deviation [sd] = 13.8), and the age distribution was: under 40 (9%; n = 28), 40–49 (13%; n = 41), 50–59 (23%; n = 72), 60–69 (32%; n = 101) and 70 and older (24%; n = 77) years. Other information, including ethnicity, parity and OC use, was not available for case subjects.

There were two independent control groups. The first consisted of 300 adult female unrelated monozygotic twins selected from a sample of 3,348 twins of almost exclusively European descent, recruited through the volunteer Australian Twin Registry for the Semi Structured Assessment for the Genetics of Alcoholism (SSAGA) research study (Heath *et al.*, 1998). Twin control subjects had participated in a telephone interview follow-up in 1992–1993, and those providing blood samples for DNA studies between 1993 and 1996 lived in or close to Adelaide, Brisbane, Melbourne or Sydney. Criteria for selection for this study were that subjects should be monozygotic (DNA from dizygotic twin pairs was in high demand for other projects), female, and that the date-of-birth distribution of the twin control subjects should match as closely as possible that of the case subjects, namely one third from each of

TABLE 1 – ANDROGEN RECEPTOR EXON 1 CAG_N GENOTYPE DISTRIBUTION IN OVARIAN CANCER CASE SUBJECTS AND CONTROL SUBJECTS[1]

i, j	nij case subjects	n*ij* twin control subjects	n*ij* population control subjects	*i, j*	nij case subjects	n*ij* twin control subjects	n*ij* population control subjects	*i, j*	nij case subjects	n*ij* twin control subjects	n*ij* population control subjects
8 17	0	0	1	17 28	0	0	2	21 29	2	1	6
8 18	1	0	0	17 29	1	0	0	21 30	0	0	2
9 20	0	1	0	17 30	0	0	1	21 31	0	0	1
9 24	0	0	1	18 18	0	1	3	21 32	0	0	2
10 23	0	1	0	18 19	3	4	8	21 37	0	0	1
11 24	0	0	1	18 20	4	4	9	22 22	4	3	10
11 26	0	0	1	18 21	1	5	10	22 23	13	6	14
12 18	0	0	1	18 22	3	2	4	22 24	11	4	14
12 21	0	1	0	18 23	4	4	10	22 25	6	4	7
12 25	0	0	1	18 24	7	5	7	22 26	5	5	6
13 13	1	0	0	18 25	1	0	3	22 27	1	1	2
13 19	0	0	1	18 26	1	7	3	22 28	0	1	0
13 20	0	0	1	18 27	1	2	4	22 29	0	1	3
13 23	1	0	1	18 28	0	0	1	22 30	1	0	0
13 24	2	1	0	18 30	0	2	0	22 31	0	0	1
13 25	0	0	1	19 19	8	2	2	23 23	4	6	9
14 18	0	0	1	19 20	9	5	17	23 24	8	2	12
14 19	0	0	2	19 21	11	9	17	23 25	5	6	2
14 20	0	1	1	19 22	9	8	9	23 26	3	5	10
14 21	1	0	0	19 23	7	7	17	23 27	1	2	6
14 23	1	2	0	19 24	4	6	14	23 28	3	1	2
14 24	0	0	1	19 25	8	4	6	23 29	0	1	1
15 18	0	0	1	19 26	4	1	4	23 30	1	1	0
15 20	0	1	0	19 27	1	2	1	23 32	0	1	0
15 21	1	1	1	19 28	0	2	3	23 35	0	0	1
15 22	0	0	1	19 29	0	1	0	23 42	0	0	1
15 23	0	0	1	19 30	1	0	1	24 24	4	3	5
15 24	0	0	1	20 20	7	3	5	24 25	5	4	15
15 27	0	0	1	20 21	9	19	19	24 26	7	5	3
16 17	1	0	0	20 22	5	8	12	24 27	0	1	4
16 19	0	0	2	20 23	7	7	13	24 28	1	0	2
16 20	1	0	2	20 24	8	11	13	24 30	1	0	0
16 22	0	0	2	20 25	4	3	7	24 32	1	0	0
16 23	1	0	1	20 26	3	6	6	25 25	2	4	1
16 24	0	0	2	20 27	3	2	5	25 26	2	3	6
16 26	1	0	0	20 28	0	3	4	25 27	1	1	3
16 27	0	0	1	20 29	2	1	0	25 28	1	2	1
17 17	0	1	0	20 30	0	1	1	25 29	0	0	1
17 18	0	0	4	21 21	14	6	16	26 26	2	2	0
17 19	1	1	1	21 22	12	6	20	26 27	0	1	0
17 20	2	1	2	21 23	15	15	26	26 28	1	2	0
17 21	0	0	1	21 24	13	14	22	27 27	0	0	1
17 22	1	1	3	21 25	7	7	13	27 28	0	0	1
17 23	1	0	2	21 26	4	8	9	32 32	1	0	0
17 24	2	0	0	21 27	2	4	3				
17 25	0	0	3	21 28	0	1	0				

[1] *i* and *j* represent the two CAG_n alleles of a given genotype; therefore, the numbers in the columns marked "*i, j*" represent the absolute number of CAG repeats in alleles *i* and *j*, respectively. nij refers to the number of subjects with the indicated "*ij*" genotypes.

1900–1925, 1926–1938 and 1939–1970. Age at interview of twin control subjects ranged from 30 to 90 years, with an average of 50.9 years (sd = 13.9), and the age distribution was: under 40 (28%; n = 85), 40–49 (20%; n = 61), 50–59 (22%; n = 65), 60–69 (19%; n = 56) and 70 and older (11%; n = 33) years. Parity data collected at interview was available for 293 of the 300 twin control subjects, and the sample included 38 (13%) nulliparous women, 170 (58%) women with 1–3 live births, and 85 (29%) women with 4 or more live births.

The second control group consisted of 553 adult Australian women without breast or ovarian cancer from the Australian Breast Cancer Family Study (McCredie *et al.*, 1998) conducted in Melbourne and Sydney from 1992–1999. These population control subjects were selected from the electoral roll (adult registration for voting is compulsory in Australia) by use of stratified random sampling. Age at interview ranged from 20 years up to 69 years, and the age distribution was: under 40 (63%; n = 351), 40–49 (16%; n = 91), 50–59 (13%; n = 73), and 60–69 (7%; n = 38) years. Parity data collected at interview was available for 492 of the 553 population control subjects, and the sample included 134 (27%) nulliparous women, 318 (65%) women with 1–3 live births, and 40 (8%) women with 4 or more live births.

Ethics clearance for collection of subject information and a blood sample from case subjects and twin control subjects was given by the Queensland Institute of Medical Research and Royal Brisbane Hospital Ethics committees, and for population control subjects by The University of Melbourne and the New South Wales Cancer Council.

Molecular analysis

Germline DNA was extracted from peripheral blood from both case subjects and twin control subjects by the salt-precipitation method (Miller *et al.*, 1988), and from population control subjects as described in Spurdle *et al.* (1999).

The androgen receptor exon 1 CAG trinucleotide repeat was amplified by polymerase chain reaction (PCR) using primer sequences detailed by La Spada *et al.* (1991), with inclusion of a 5′-6-carboxy-4,7,2′,7′-tetrachlorofluoroscein(TET)-labeled forward primer to generate a fluorescent product. The 10 μL reaction mix contained 30 ng DNA, primers (10 pmol each), deoxynucle-

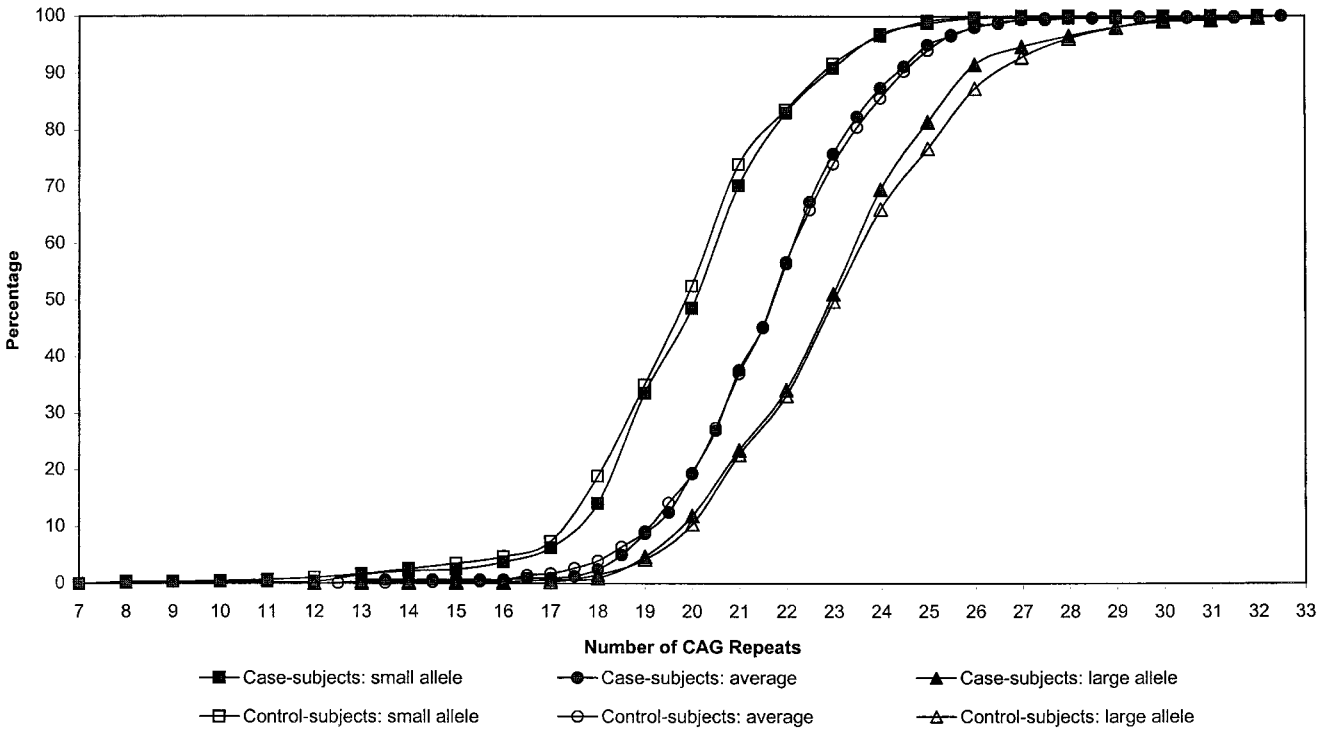

FIGURE 2 – Androgen receptor exon 1 CAG_n genotype-allele cumulative frequency distribution in ovarian cancer case subjects and pooled twin and population control subjects. The terms smaller, larger and average refer to the smaller, larger and arithmetic mean of the two alleles of an individual's genotype.

otidetriphosphates (200 nM), 1× Perkin-Elmer Taq polymerase buffer, 1 U Taq polymerase, 1.5 mM $MgCl_2$ and 7% deionized formamide. Amplification conditions were 2 min at 94°C, and 34 cycles at 94°C for 20 s, 62°C for 20 s and 72°C for 20 s, followed by a 10 min extension at 72°C. Amplified samples were diluted one in 12 in formamide loading buffer, denatured for 2 min at 95°C, and size-separated on a 6% denaturing polyacrylamide gel. The ABI Prism 373 Genescan and Genotyper systems (Perkin-Elmer Corp., Foster City, CA) were used for detection and sizing of fluorescent products. Separation of the ABI TAMRA-350 size standard in each lane allowed for Genescan automated sizing of 5'-TET-labelled PCR products. In addition, control samples of known size were separated at different positions across each gel. Consistent sizing of samples was indicated by the independent generation of matching size results for a random subset (17.9%) of samples separated on more than one gel. PCR-amplified samples from both case subjects and control subjects were loaded randomly on gels to further avoid any sizing bias.

Statistical Analysis

Case subjects and control subjects were compared for CAG_n firstly as a continuum. Differences in mean between case and control groups were tested using Student's *t*-test, and analysis of covariance to allow adjustment for any age effects. Linear regression was used to test for associations with age within case and control groups, and associations with parity within control groups. For CAG_n dichotomised by a cut-point, crude comparisons were made using standard tests of proportion and the χ^2 distribution. Unconditional logistic regression with adjustment for age was used to estimate risk of ovarian cancer in terms of an odds ratio (OR) and 95% confidence interval (CI). We used a cut-point of ≥ 22, because this divided the distribution of CAG_n approximately in half, was near to the observed mode of 21, and was a cut-point reported to show an association with prostate cancer (Stanford *et al.*, 1997). We also used the cut-point ≥ 29, as this has been implicated in modifying breast cancer risk in BRCA1 carriers (Rebbeck *et al.*, 1999), and the cut-point $CAG_n \geq 27$, since inspection of the data (Fig. 1) suggested that control subjects had a greater frequency of this allele relative to case subjects. For genotypes defined by the cut-point ≥ 22 CAG_n, the linear-by-linear association test was used to assess differences in the distribution of CAG_n with increasing parity.

Estimation and comparison of allele frequencies, and the standard test of Hardy-Weinberg Equilibrium, were undertaken by maximum likelihood methods and reference to the χ^2 distribution, as detailed in Spurdle *et al.* (1999). Post-hoc power calculations were based on using the observed standard errors to derive the standard error of the test statistic (difference in means or effect of genotype on log odds of being a case subject) and multiplying by $0.84 + 1.96 = 2.80$, for 80% power at 0.05 two-tailed. Statistical analyses were performed using the software packages SPSS, Egret and Epi-Info. All *p* values are two-tailed and, following convention, statistical significance is taken as a nominal p value < 0.05.

RESULTS

Genotype data were generated for 319 cases and a total of 853 controls (Table I). The allele frequency distributions of case and control subjects are shown in Figure 1, with a mode of 21 CAG_n, and a median of 22 CAG_n for all three groups. The mean (95% CI) CAG_n was 21.9 (21.7–22.1) for case subjects, 22.1 (21.9–22.3) for twin control subjects and 21.9 (21.7–22.1) for population control subjects. There was no difference in mean CAG_n between the two control subject groups ($p = 0.2$), or between case subjects and control subjects ($p = 0.4$, $p = 0.7$ and $p > 0.9$ for twin control, population control and pooled control subject groups respectively).

Figure 2 shows the cumulative distribution curves for the smaller, larger and average CAG_n of the CAG_n genotype, for case subjects and the pooled control subjects. There were no differences between the two control groups for the means of the smaller, larger and average CAG_n ($p = 0.2$, 0.3 and 0.2, respectively). Pooling

control subjects, there were no differences in means between case subjects and control subjects for the smaller, larger and average CAG_n. The mean (95% CI) for the smaller CAG_n was 20.5 (20.2–20.8) for case subjects and 20.3 (20.1–20.5) for control subjects ($p = 0.2$). For the larger CAG_n, these means were 23.4 (23.1–23.7) and 23.6 (23.4–23.8), respectively ($p = 0.2$), while for the average CAG_n they were 22.0 (21.8–22.2) and 22.0 (21.9–22.1) respectively ($p = 0.95$). There was 80% power at the 0.05 level of significance to detect differences in means in excess of 0.2 standard deviations, or 0.6 repeat length.

Comparing the subgroup of case subjects with invasive ovarian cancer to the pooled control subjects, the mean (95% CI) for invasive cancer case subjects was 20.5 (20.3–20.9), 23.3 (23.0–23.6) and 22.0 (21.8–22.2) for the smaller, larger or average CAG_n, respectively. There was no plausible evidence for a difference between invasive cancer case subjects and control subjects, since although there was a suggestion that invasive cancer case subjects had a greater mean for the smaller allele ($p = 0.05$), the shift in size was *reversed* for the mean for the larger allele ($p = 0.09$), and there was no difference for the average allele size ($p = 0.99$).

Although mean age of case subjects was older than that of either control subject group, linear regression suggested that there was no significant association between age and the smaller, larger or average CAG_n in case subjects ($p = 0.07$, 0.3, and 0.1, respectively) or population control subjects ($p = 0.3$, 0.8 and 0.6), while there was a suggestion that the mean for the smaller, larger and average CAG_n decreased with increasing age in twin control subjects ($p = 0.02$, 0.05 and 0.01). However, there were no differences in mean CAG_n when case subjects were compared with pooled control subjects after allowing for age ($p = 0.2$, 0.3 and 0.9 for smaller, larger and average CAG_n, respectively). Epidemiological information such as parity was unavailable for case subjects, and it was thus not possible to compare mean CAG_n allowing for the effects of parity and other known ovarian cancer risk factors. However, there was no indication that differences in the distribution of CAG_n could have been masked by differences in parity between case and control subjects, since there were no differences in mean CAG_n with increasing parity in either control group ($p > 0.5$ for all analyses).

Table II shows the genotype frequency data for the cut-points ≥ 22 and ≥ 29, and includes both crude and age-adjusted estimates of OR for the association between genotype and risk of ovarian cancer. There was no evidence of an association between genotype and risk of ovarian cancer for either cut-point, either before or after adjustment for age. Again, although it was not possible to adjust for parity, there were no differences in the distribution of CAG_n with parity in either control group, for the cut-points ≥ 22 and ≥ 29 CAG_n ($p > 0.3$ for all analyses). We had 80% power at 0.05 to detect ORs of 1.5 or more, or 0.63 or less.

In addition to the above-mentioned cut-points, the cut-point $CAG_n \geq 27$ was evaluated with respect to ovarian cancer risk, since inspection of the data in Figure 1 suggested that control subjects had a greater frequency of this allele relative to case subjects. Comparing genotype frequencies as for the other cut-points, there was a suggestion that alleles ≥ 27 were protective for ovarian cancer, since unadjusted comparisons yielded an OR of 0.64, with 95% CIs of 0.38–1.07 ($p = 0.09$), 0.40–1.02 ($p = 0.06$) and 0.41–0.99 ($p = 0.05$) for twin control, population control and pooled control comparisons respectively. Results differed little after adjustment for age, with ORs ranging from 0.52 (comparing to population control-subjects) to 0.64 (comparing to pooled control subjects), but now the 95% confidence intervals excluded unity for the twin control analysis alone. There was no deviation from Hardy-Weinberg equilibrium in either case subjects or pooled control subjects for any of the cut-points ($p = 0.3$–0.8), with the exception of the cut-point ≥ 29 CAG_n in case subjects ($p = 0.007$).

TABLE II – ASSOCIATION OF AR CAG_N GENOTYPE STATUS WITH OVARIAN CANCER

CAG_n Cut-point	Number of alleles	Case subjects (n)	Twin control subjects (n)	Population control subjects (n)	Comparison to twin control subjects				Comparison to population control subjects				Comparison to pooled control subjects			
					Crude OR	[95% CI]	Age-adjusted OR	[95% CI]	Crude OR	[95% CI]	Age-adjusted OR	[95% CI]	Crude OR	[95% CI]	Age-adjusted OR	[95% CI]
≥22	0	75	66	128	Reference		Reference		Reference		Reference		Reference		Reference	
	1	149	156	281	0.84	[0.55–1.28]	0.88	[0.58–1.34]	0.90	[0.64–1.28]	0.75	[0.48–1.18]	0.88	[0.64–1.22]	0.86	[0.59–1.25]
	2	95	78	144	1.07	[0.67–1.72]	1.16	[0.72–1.85]	1.13	[0.77–1.65]	1.01	[0.61–1.67]	1.11	[0.77–1.58]	1.18	[0.78–1.78]
	1 or 2	244	234	425	0.92	[0.62–1.36]	0.97	[0.65–1.44]	0.98	[0.71–1.36]	0.84	[0.55–1.28]	0.96	[0.71–1.30]	0.96	[0.68–1.37]
≥29	0	308	290	530	Reference		Reference				Reference				Reference	
	1 or 2	11	10	23	1.04	[0.43–2.48]	1.01	[0.40–2.53]	0.82	[0.40–1.71]	1.02	[0.38–2.74]	0.89	[0.44–1.78]	1.06	[0.47–2.38]

Case subjects were also stratified by cancer form, histology, grade and stage to investigate the possibility of AR genotype heterogeneity between ovarian cancer subgroups for the cut-point $\geq$ 22 CAG_n. There were no major differences in genotype distribution between any of the subgroups ($p = 0.9$ for cancer form, 0.7 for histology, 0.13 for stage, and 0.7 for grade).

DISCUSSION

The data presented in this Australian study do not provide any convincing evidence for an association between AR exon 1 CAG repeat length and ovarian cancer risk. We cannot, however, exclude small effects, such as a difference in mean CAG_n of 0.2 standard deviations or less, or an effect of less than 1.5 for CAG_n above or below the median cut-point. We also cannot exclude threshold effects in a small subgroup. Rebbeck *et al.* (1999) reported that alleles $\geq$ 29 CAG_n were associated with earlier age at onset of breast cancer in women who carried a mutation in BRCA1. The BRCA1 mutation status of our subjects has yet to be determined. In our study, there was no evidence for an increased risk of ovarian cancer for alleles $\geq$ 29 CAG_n, but only 10 case subjects and 23 control subjects had an allele this large, so there was little power to detect effects unless they were 3-fold or more.

Although there was a suggestion that alleles $\geq$ 27 CAG_n may be weakly protective for ovarian cancer, this was of borderline statistical significance. Given that comparisons were carried out for a number of cut-points, this marginally significant effect should be interpreted with caution. Such a protective effect, however, would be consistent with the hypothesis of Risch (1998) that increased androgen exposure is a risk factor for ovarian cancer, because functional studies (Chamberlain *et al.,* 1994) have suggested that larger CAG_n alleles are associated with decreased hormone action of the AR.

We stratified the ovarian cancer cases according to several clinical criteria to assess whether there was heterogeneity with respect to AR genotype, because there is evidence to suggest that the mutational pathway differs between invasive tumours and benign or low malignant potential ovarian tumours (Berchuck *et al.*, 1994; Chenevix-Trench *et al.*, 1997), and between tumours of different histology (Kvale *et al.*, 1988; Ichikawa *et al.*, 1994; Fujita *et al.*, 1995; Obata *et al.*, 1998). Stratification of the ovarian cancer sample according to form, histology, grade or stage failed to reveal any heterogeneity with respect to AR genotype defined by the median CAG_n cut-point of 22.

There was no reason to suspect that selection of younger control groups had influenced the findings, since there was no consistent evidence that the allele size distribution depended on age in our case or control subjects, and there were no marked differences between crude and age-adjusted ORs for the different CAG_n cut-points. Although population differences in CAG_n distribution have been reported (Edwards *et al.*, 1992, Irvine *et al.*, 1995), confounding due to differences in ethnicity was also improbable, since the great majority of Australians are Anglo-Celtic in origin, and case subjects and control subjects were unlikely to differ appreciably with respect to their ethnic background.

In conclusion, the AR exon 1 CAG repeat polymorphism does not appear to influence ovarian cancer risk. We are conscious that analysis of data at the extremes of the allele range is limited in power, and emphasize that replication is essential in establishing credible results. As in Spurdle *et al.* (1999), we thus present the data in their raw form to allow pooling with other similar population-based studies, and encourage others to do likewise.

ACKNOWLEDGEMENTS

We thank Soo-Keat Khoo, Terry Hurst and Bruce Ward of the Department of Obstetrics and Gynaecology, The University of Queensland, for collection of ovarian case subject material. We are grateful to Judith Kerr from the Queensland Institute of Medical Research (QIMR) for preparation of DNA from ovarian case subjects, Melissa Southey and Hun Chy from the Peter MacCallum Cancer Institute for preparation of DNA from population control subjects, John Pearson from the QIMR for access to and selection of twin control subject DNA and Carol Mayne, Alana Goldman and Joanne Voisey for technical assistance with this project. Confirmation of results was possible through the provision of size standards by Steve Edwards of the Institute of Cancer Research, Sutton, Surrey, Najah Nassif of Sydney University, and Wayne Tilley and Grant Buchanan of Flinders University. We also thank Margaret R. E. McCredie of the NSW Cancer Council and Beth Newman of the Queensland Institute of Technology for helpful discussion and comments on this manuscript. We are grateful to the many women who participated in this research, and to the Australian Twin Registry. This work was supported by funds from the National Health and Medical Research Council of Australia, the Victorian Health Promotion Foundation, the NSW Cancer Council, and the Peter MacCallum Cancer Institute.

REFERENCES

AL-TIMIMI A., BUCKLEY C.H. and FOX H., An immunohistochemical study of the incidence and significance of sex steroid hormone binding sites in normal and neoplastic human ovarian tissue. *Int. J. Gynecol. Pathol.,* **4**, 24–41 (1985).

AUSTRALIAN INSTITUTE OF HEALTH AND WELFARE (AIHW) AND AUSTRALASIAN ASSOCIATION OF CANCER REGISTRIES (AACR). *Cancer in Australia 1995: Incidence and mortality data for 1995 and selected data for 1996.* AIHW cat. No. CAN 5. (Cancer Series No. 10), Australian Institute of Health and Welfare, Canberra (1998).

BERCHUCK, A., KOHLER, M.F., MARKS, J.R., WISEMAN, R., BOYD, J. and BAST, R.C., The p53 tumor suppressor gene frequently is altered in gynecologic cancers. *Am. J. Obstet. Gynecol.,* **170,** 246–252 (1994).

BOYD, J. and RUBIN S.C., Hereditary ovarian cancer: molecular genetics and clinical implications. *Gyn. Oncol.,* **64**, 196–206 (1997).

CHAMBERLAIN, N.L., DRIVER, E.D. and MIESFIELD, R.L., The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucl. Acids Res.,* **22**, 3181–3186 (1994).

CHENEVIX-TRENCH, G., KERR, J., HURST, T., SHIH, Y-C., PURDIE, D., BERGMAN, L., FRIEDLANDER, M., SANDERSON, B., ZOURNAZI, A., COOMBS, T., LEARY, J.A., CRAWFORD, E., SHELLING, A.N., COOKE, I., GANESAN, T.S., SEARLE, J., CHOI, C., BARRETT, J.C., KHOO, S-K. and WARD, B., Analysis of loss of heterozygosity and KRAS2 mutations in ovarian neoplasms: clinicopathological correlations. *Genes Chrom. Cancer,* **18**, 75–83 (1997).

EDWARDS, A., HAMMOND, H.A., JIN, L., CASKEY, C.T. and CHAKRABORTY, R., Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics,* **12**, 241–253 (1992).

FUJITA, M., ENOMOTO, T., YOSHINO, K., NOMURA, T., BUZARD, G.S., INOUE, M. and OKUDAIRA Y., Microsatellite instability and alterations in the hMSH2 gene in human ovarian cancer. *Int. J. Cancer,* **64**, 361–366 (1995).

GAYTHER, S.A., RUSSELL, P., HARRINGTON, P., ANTONIOU, A.C., EASTON, D.F. and PONDER, B.A.J., The contribution of germline BRCA1 and BRCA2 mutations to familial ovarian cancer: no evidence for other ovarian cancer-susceptibility genes. *Am. J. Hum. Genet.,* **65**, 1021–1029 (1999).

GIOVANUCCI, E., STAMPFER, M.J., KRITHIVAS, K., BROWN, M., BRIFSKY, A., TALCOTT, J., HENNEKENS, C.H. and KANTOFF, P.W., The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc. Natl. Acad Sci. USA,* **94**, 3320–3323 (1997).

GIOVANNUCCI, E., PLATZ, E.A., STAMPFER, M.J., CHAN, A., KRITHIVAS, K., KAWACHI, I., WILLETT, W.C. and KANTOFF, P.W., The CAG repeat within the androgen receptor gene and benign prostatic hyperplasia. *Urology,* **53**, 121–125 (1999).

HAKIMI, J.M., SCHOENBERG, M.P., RONDINELLI, R.H., PIANTADOSI, S. and BARRACK, E.R., Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clin. Cancer. Res.,* **3**, 1599–1608 (1997).

HARDY, D.O., SCHER, H.I., BOGENREIDER, T., SABBATINI, P., ZHANG, Z.F., NANUS, D.M. and CATTERALL, J.F., Androgen receptor CAG repeat lengths in prostate cancer: correlation with age at onset. *J. Clin. Endocrinol. Metab.,* **81**, 4400–4405 (1996).

HEATH, A.C., BUCHOLZ, K.K., MADDEN, P.A., DINWIDDIE, S.H., SLUTSKE, W.S., BIERUT, L.J., STATHAM, D.J., DUNNE, M.P., WHITFIELD, J.B. and MARTIN, N.G., Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychol. Med.*, **27**, 1381–1396 (1998).

HELZLSOUER, K.J., ALBERG, A.J., GORDON, G.B., LONGCOPE, C., BUSH, T.L., HOFFMAN, S.C. and COMSTOCK, G.W., Serum gonadotropins and steroid hormones and the development of ovarian cancer. *JAMA*, **274**, 1926–1930 (1995).

ICHIKAWA, Y., NISHIDA, M., SUZUKI, H., YOSHIDA, S., TSUNODA, H., KUBO, T., UCHIDA, H. and MIWA, M., Mutation of K-ras protooncogene is associated with histological subtypes in human mucinous ovarian tumours. *Cancer Res.*, **54**, 33–35 (1994).

INGLES, S.A., ROSS, R.K., YU, M.C., IRVINE, R.A., LA PERA, G., HAILE, R.W. and COETZEE, G.A., Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. *J. Natl. Cancer. Inst.*, **89**, 166–170 (1997).

IRVINE, R.A., YU, M.C., ROSS, R.K. and COETZEE, G.A., The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Res.*, **55**, 1937–1940 (1995).

KVALE, G., HEUCH, I., NILSSON, S. and BERAL, V., Reproductive factors and risk of ovarian cancer: a prospective study. *Int. J. Cancer*, **42**, 246–251 (1988).

LA SPADA, A.R., WILSON, E.M., LUBAHN, D.B., HARDING, A.E. and FISCHBECK, K.H., Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*, **352**, 77–79 (1991).

MACLEAN, H.E., CHOI, W.T., REKARIS, G., WARNE, G.L. and ZAJAC, J.D., Abnormal androgen receptor binding affinity in subjects with Kennedy's disease (spinal and bulbar muscular atrophy). *J. Clin. Endocrinol. Metab.*, **80**, 508–516 (1995).

MCCREDIE, M.R., DITE, G., GILES, G.G. and HOPPER, J.L., Breast cancer in Australian women under the age of 40. *Cancer Cause. Control*, **9**, 189–198 (1998).

MILLER, S.A., DYKES, D.D. and POLESKY, H.F., A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl. Acids Res.*, **16**, 1215 (1988).

OBATA, K., MORLAND, S.J., WATSON, R.H., HITCHCOCK, A., CHENEVIX-TRENCH, G., THOMAS, E.J. and CAMPBELL, I.G., Frequent PTEN/MMAC mutations in endometriod but not serous or mucinous epithelial ovarian tumours. *Cancer Res*., **58**, 2095–2020 (1998).

PARAZZINI, F., FRANCESCHI, S., LA VECCHIA, C. and FASOLI, M., The epidemiology of ovarian cancer. *Gynecol. Oncol*., **43**, 9–23 (1991).

PETTERSSON, F., *Annual report on gynaecologic cancer of FIGO* (Vol 20). Panorama Press, Stockholm (1988).

PURDIE, D., GREEN, A., BAIN, C., SISKIND, V., WARD, B., HACKER, N., QUNII, M., WRIGHT, G., RUSSELL, P. and SUSIL, B., Reproductive and other factors and risk of epithelial ovarian cancer: an Australian case-control study. Survey of Women's Health Study Group. *Int. J. Cancer*, **15**, 678–684 (1995).

REBBECK, T.R., KANTOFF, P.W., KRITHIVAS, K., NEUHAUSEN, S., BLACKWOOD, M.A., GODWIN, A.K., DALY, M.B., NAROD, S.A., GARBER, J.E., LYNCH, H.T., WEBER, B.L. and BROWN, M., Modification of BRCA1-associated breast cancer risk by the polymorphic androgen receptor CAG repeat. *Am. J. Hum. Genet.*, **64**, 1371–1377 (1999).

RISCH, H.A., Hormonal Etiology of epithelial ovarian cancer, with a hypothesis concerning the role of androgens and progesterone. *J. Natl. Cancer. Inst.*, **90**, 1774–1786 (1998).

SILVA, E.G., TORNOS, C., FRITSCHE, H.A. JR, EL-NAGGAR, A., GRAY, K., ORDONEZ, N.G., LUNA, M. and GERSHENSON, D., The induction of benign epithelial neoplasms of the ovaries of guinea pigs by testosterone stimulation: a potential animal model. *Mod. Pathol.*, **10**, 879–883 (1997).

SPURDLE, A.B., DITE, G.S., CHEN, X., MAYNE, C.J., SOUTHEY, M.C., BATTEN, L., CHY, H., TRUTE, L., MCCREDIE, M.R.E., GILES, G.G., ARMES, J., VENTER, D.J., HOPPER, J.L. and CHENEVIX-TRENCH, G., Androgen Receptor Exon 1 CAG repeat length and breast cancer in women before age forty years. *J. Natl. Cancer Inst.*, **91**, 961–966 (1999).

STANFORD, J.L., JUST, J.J., GIBBS, M., WICKLUND, K.G., NEAL, C.L., BLUMENSTEIN, B.A. and OSTRANDER, E.A., Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res.*, **57**, 1194–1198 (1997).

STRATTON, J.F., GAYTHER, S.A., RUSSEL, P., DEARDEN, J., GORE, M., BLAKE, P., EASTON, D. and PONDER, B.A.J., Contribution of BRCA1 mutations to ovarian cancer. *New Engl. J. Med.*, **336**, 1125–1130 (1997).

TUT, T.G., GHADESSY, F.J., TRIFIRO, M.A., PINSKY, L. and YONG, E.L., Long polyglutamine tracts in the androgen receptor are associated with reduced transactivation, impaired sperm production, and male infertility. *J Clin. Endocrinol. Metab.*, **82**, 3777–3782 (1997).