

SHORT REPORT

Cervix smear abnormalities: linking pathology data in female twins, their mothers and sisters

Jacqueline M Vink^{*,1}, Folkert J van Kemenade², Chris JLM Meijer², Mariel K Casparie³, Gerrit A Meijer² and Dorret I Boomsma¹

Mass screening for cervical cancer precursors has decreased the incidence of cervical cancer in several countries, including the Netherlands. Persistent infections of certain types of human papillomavirus are strongly associated with the development of cervical cancer. A number of factors may affect the liability to infection and subsequent progression to cervical intraepithelial neoplasia and cancer. This paper examines whether genetic factors are involved in explaining individual differences in liability. Data of 3178 women registered with the Netherlands Twin Register were successfully linked to the nationwide Dutch Pathology database that contains all results of mass screening for cervical cancer. The data from mono- and dizygotic twins and their female relatives were used to disentangle the influence of heritable and environmental factors on cervix smear abnormalities. Results showed that differences in cervix smear results clustered within families and resemblance was stronger in monozygotic twins (correlation 0.37, 95% confidence interval: 0.12–0.58) compared with other first-degree relatives (correlation 0.14, 95% confidence interval: –0.01–0.29). The familial clustering for an abnormal cervix smear is due to shared genetic factors that explain 37% of the variance in liability. The largest proportion of the variation in cervical smear abnormalities is due to unique environmental factors.

European Journal of Human Genetics (2011) 19, 108–111; doi:10.1038/ejhg.2010.139; published online 18 August 2010

Keywords: cervix abnormalities; twins; cervical cancer

INTRODUCTION

Mass screening for cervical cancer has been introduced in several countries to decrease the incidence of cervical cancer.^{1,2} Cytological screening of cervix smears can detect neoplastic changes of the cervical epithelium. Cervical cancer typically has a long preinvasive state (often a decade or more) and the treatment for preinvasive disease is effective.³ In the Netherlands, screening was introduced in 1988 and cervical cancer incidence decreased from 12–18 per 100 000 women-years in the period 1960–1970 to 7.5 women per 100 000 women-years in the past years.^{4,5}

Several studies observed familial aggregation for cervical cancer and risk ratios between 1.8 and 3.2 have been reported in first-degree relatives.^{6–10} A seminal study in Scandinavian twins suggested that the familial aggregation for cervical cancer is not caused by shared genes, but should be attributed to shared environment because the concordance in monozygotic twins (MZ) (genetically identical) was not higher than the concordance in dizygotic twin (DZ) pairs (on average 50% shared genes).¹¹ In contrast, a Swedish twin study reported a moderate heritable effect of *in situ* cases.¹²

In the present study we explore an intermediate endpoint for cervix cancer, for example, cervical intraepithelial neoplasia (CIN) measured by cervix smear abnormalities. Data of 3178 women registered with the Netherlands Twin Register were successfully linked to the nationwide Dutch Pathology database (PALGA) that includes results of cervix smears of Dutch women.^{13,14} We investigate the familial resemblance of cervix smear phenotype in twins, their sisters and

their mothers and use the twin/family data to disentangle the influence of heritable and environmental factors on cervix smear abnormalities.

METHODS

Sample

Since 1991 nationwide coded data in the Netherlands on all pathology diagnoses made in Dutch pathology laboratories, including cytology results, are stored in a central PALGA, both for clinical and research purposes.¹⁴ This includes structured data on the nationwide screening program for cancer of the uterine cervix. A total of 80% of Dutch women aged 30–60 years participate in this program and are screened at 5-year intervals.^{15,16}

The Netherlands Twin Register collects data of adult twins and their family members in a longitudinal study to health, lifestyle and personality.¹⁷ In the 2002 and 2004 survey participants were asked for permission to link their data to other databases in The Netherlands. In total, 6117 women were born before 1978 and were thus eligible for screening. Of this group, 5668 (92.7%) had given permission to link their data to other databases. As no unique identifier like a social security number was available, data were linked on last name (maiden name), date of birth and first initial (if available). For 292 women information on their last name/maiden name was missing. For 4110 women we were certain of their maiden name and for 1266 it was not sure whether the last name in our register was their maiden name or the last name of the spouse. Twin pairs with the same initials (and obviously same maiden name and same date of birth) were excluded ($n=66$ pairs) because it was not possible to distinguish their PALGA records. Linking the Netherlands Twin Register data to the PALGA database resulted in data on cervix smear for 3178 women (74.3% of 4110 women with known maiden name and 8.3% of remaining

¹Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands; ²Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands; ³Foundation PALGA, Utrecht, The Netherlands

*Correspondence: Dr JM Vink, Department of Biological Psychology, VU University Amsterdam, Van der Boechorststraat 1, Amsterdam 1081 BT, The Netherlands. Tel: +31 20 598 8787; Fax: +31 20 598 8832; E-mail: jm.vink@psy.vu.nl

Received 22 March 2010; revised 22 June 2010; accepted 29 June 2010; published online 18 August 2010

1266 women). This sample consisted of 2020 twins, 671 sisters of twins, 416 mothers of twins and 71 female spouses of male twins. The twins included 1045 MZ, 588 dizygotic same-sex twins and 170 women from dizygotic opposite-sex pairs. Zygosity was unknown for 30 women (including 16 triplet members). Year of birth ranged from 1919 to 1977 (median: 1968). Mean age at last screening was 38.1 years (SD 9.2) for twins and sisters and 53.0 years (SD 7.7) for the mothers of twins.

Phenotype

The cervical smears were classified into three categories: no abnormalities, borderline/mild dyskaryosis (comparable with atypical cells of undetermined significance: ASCUS/low-grade squamous intraepithelial lesions: in the Bethesda 2001 classification) or moderate dyskaryosis or worse (high-grade squamous intraepithelial lesions: or worse). For each woman, the highest score ever was analyzed. When cervical smear results were of insufficient quality they were excluded. In 1996, new guidelines for the classification and management of cervical smears in the Dutch population screening program for cervical cancer were implemented,^{15,18} and as a consequence the prevalence of cervical smear abnormalities changed. We adjusted for this in the statistical analyses (see below).

Statistical analyses and study design

A liability threshold-model was used to explore familial resemblances.¹⁹ This model assumes an underlying (latent) liability to an ordinal variable (cervical smear results). This liability is the sum of the effects of many genetic and environmental factors and has a normal distribution and unit variance. Two thresholds divide the sample into three groups: no abnormalities, borderline/mild abnormalities or moderate/severe abnormalities. The thresholds are based on the observed prevalence and can be interpreted as *z*-score.

Thresholds were allowed to vary as a function of age at last screening and the year of diagnosis. Age at last screening and the year of diagnosis were included as dichotomies with <40 years versus ≥40 years for age at last screening. For women with a diagnosis (one or two) we distinguished between a diagnosis before 1996 and in 1996 or later (in 1996 there was a change in the diagnostic system). For women who screened negative (diagnosis 0) year of diagnosis was defined on the basis of whether their last screening was before 1996 or not.

Quantitative genetic analyses were carried out using the software package Mx,²⁰ and models were fitted to the raw data. Familial resemblance for liability to cervical smear abnormality was expressed as polychoric correlations for MZ twins and first-degree relatives (DZ twins, sister and mother–daughter pairs). The polychoric correlation coefficient is a measure of association between two ordinal variables. In this study the ordinal variable consists of cervix smear abnormalities in three categories. MZ pairs are genetically identical, whereas other first-degree relatives, such as DZ twins and siblings, share on average 50% of their segregating genes. Each parent transmits 50% of his/her genes to a child, so mother and daughter pairs share 50% of their genes. Consequently, if the MZ correlation is larger than the correlation in other first-degree relatives, we infer that genetic influences have a role. If the correlations are equal (and larger than 0), the similarity between family members is not explained by genetic, but by shared environmental factors.

Sources of variation that were considered in genetic modeling were: additive genetic factors A (heritability), common environmental factors C (shared by family members) and unique environmental factors E. Under the full model, both genetic and shared environmental factors contribute to the resemblance

between family members (model 1). The correlations between the genetic factors were 1.0 for MZ twins (genetically identical) and 0.5 for other first-degree relatives (who share on average 50% of their segregating genes). The correlation for the shared environmental effects was set to 1.0 for all family relations.^{19,21} In model 2, both genetic factors (A) and shared environmental factors (C) are dropped simultaneously to test whether the familial clustering is significant. In model 3, the genetic factor (A) is excluded, whereas in model 4, the shared environmental factor (C) is dropped. The significance of the parameters that differ between nested models can be examined using the difference in $-2 \times \log$ likelihood that is distributed as a χ^2 statistic (with the same degrees of freedom as the difference in the degrees of freedom between the two models). With the χ^2 statistic and difference in degrees of freedom a *P*-value is obtained indicating whether the sub-model fits significantly worse (if $P < 0.05$). The best model is a model with the smallest number of parameters necessary to explain the data. The Akaike's Information Criterion is a measure of the parsimony of the model and a lower value of Akaike's Information Criterion indicates a more parsimonious model.

RESULTS

The cytological scores showed 2530 (79.6%) of women without abnormalities, 535 (16.9%) of women with borderline/mild abnormalities, 96 (3.0%) of women with moderate/severe abnormalities and 16 (0.5%) of women with insufficient quality of cervical smear ($n=6$) or missing results ($n=10$). The mean number of observations on cervix smear per women was 3.4 (SD 2.8, median 3). No differences in prevalence between twins and their sisters were observed.

In 1996 the scoring system changed and this influenced the prevalence of mild and severe abnormalities. Most of the women who were never classified with abnormalities had their last check after 1996 (95%). Of the 535 women with mild abnormalities, 345 (64.5%) were diagnosed before 1996, 141 (26.4%) in 1996 or later and the remaining 49 (9.1%) had a diagnosis with mild abnormalities before and after 1996. Of the 96 women who were diagnosed with moderate or severe abnormalities, 33 (34.4%) were diagnosed before 1996, 62 (64.6%) were diagnosed in 1996 or later and 1 both before and after 1996.

The prevalence of borderline/mild abnormalities and moderate/severe dyskaryosis or worse is shown in Table 1 split for two age groups. Prevalence of abnormalities was higher in the women who were diagnosed before 1996 and in the women aged above 40 years.

The correlations for MZ twin pairs and for first-degree relatives are shown in Table 2. MZ twin pairs showed a stronger resemblance than first-degree relatives indicating that genetic factors have a role in the individual differences in liability to cervix smear abnormalities.

This observation was confirmed by the model-fitting results. The full model included additive genetic factors (A), shared environmental factors (C) and unique environmental influences (E) and showed the variance explained by genetic factors was 37%, whereas the variance explained by shared environmental factors was estimated at almost 0%. The remaining variance (63%) was explained by unique environmental factors. Dropping both A and C from the model gave a

Table 1 Prevalences of cervical smear results in three categories: no abnormalities, mild abnormalities and severe abnormalities (worst abnormality on any smear) for two age groups and split for a diagnosis before 1996 or later

	< 40 years		≥ 40 years		Total
	< 1996	≥ 1996	< 1996	≥ 1996	
No abnormalities	61 (51.3%)	1642 (92.3%)	69 (17.7%)	758 (86.7%)	2530 (80.0%)
Borderline/mild dyskaryosis	53 (44.5%)	91 (5.1%)	292 (75.1%)	99 (11.3%)	535 (16.9%)
Moderate/severe dyskaryosis or worse	5 (4.2%)	46 (2.6%)	28 (7.2%)	17 (1.9%)	96 (3.0%)
Total	119	1779	389	874	3161

significant reduction in the goodness of fit of the model (model 2), showing that familial clustering is significant. The statistical power of the analyses made it difficult to distinguish between the CE (model 3) and AE (model 4) model, compared with the full ACE model. The Akaike's Information Criterion indicated that model 4, including genetic (A) and unique environmental factors (E), was the most parsimonious model (Table 3).

Women with moderate/severe dyskaryosis or worse abnormalities are sent for kolposcopy-directed biopsy. Records with histological evaluation of biopsies were available for 302 women. In total, 86 women were diagnosed with CIN 2 or 3 or carcinoma. The number of histological scores was too small to estimate family correlations. The group of 86 women with a CIN2+ score consisted mostly of unrelated individuals, except for two sister pairs both consisting of a twin and a singleton sister.

The total sample included 526 complete twin pairs (information on cervix smear was available for both members of a twin pairs). Of the 348 MZ twin pairs 330 pairs (94.8%) were concordant for no CIN2+ score (and both at least one screening of cervix smears), whereas 18 (5.2%) were discordant for CIN2+ score (one had a CIN2+ score and the other not, although both were screened for cervix smear at least once). None of the MZ pairs was concordant for CIN2+. The pattern was the same for the 178 DZ twin pairs: 169 concordant negative (94.9%), 9 discordant for CIN2+ (5.1%) and no pairs concordant for CIN2+.

DISCUSSION

Data on the contribution of inherited and environmental factors to the causation of cervix cancer in studies of twins are rare,¹¹ possibly because only a few twin registers go back far enough in time to provide enough cases of cancer for reliable conclusions to be drawn.

Instead of cervical cancer we used an intermediate endpoint, that is, CIN, measured by cervix smear abnormalities as recorded in the Dutch national pathology registry PALGA in a sample of twins and their family members.

The current paper showed significant familial clustering for cervical smear results. The similarity in monozygotic twin pairs was higher than in first-degree relatives (DZs, sisters and mothers of twins). Genetic factors influenced 37% of the variation in cervix smear

abnormalities. The remaining variance was explained by unique environmental factors. As far as the authors know, the heritability of cervix smear abnormalities has not been studied in a large-scale twin study before.

Recently, a study among women who had not responded to invitations to the regular cervical screening program was invited to submit a self-collected cervicovaginal sample for human papillomavirus (HPV) testing. Self-sampling responders who had not participated in the previous round of screening had increased relative risks compared with self-sampling women who had been screened in the previous round.²² Including data from these women might slightly change heritability estimates.

Women with abnormal cervix smear results have an increased risk to develop cervix cancer. Depending on the results of cytology they are advised to repeat the cervical smear or they are sent for colposcopy-directed biopsy because they have an increased chance to harbor a precursor lesion (CIN2+). Not all women with abnormal cervical smear results have precursor lesions and it is estimated that about 30% of CIN3 lesions (50% of persistent CIN3 lesions) will develop into carcinomas.²³ In the present study, 86 women were diagnosed with a CIN2+ score, including two sister pairs, but no twin pairs. The sample was too small to estimate the heritability.

Several studies investigated familial aggregation of cervix cancers in samples of first-degree relatives (no twins).^{6–10} In a large Swedish study the correlations for cervical tumor diagnosis in first-degree relatives ranged from 0.13 to 0.17.⁹ This study included data of biological and adoptive mothers, full, half and adoptive sisters, and the heritability estimate for cervical tumors was 27%.⁹ Another Swedish study using family data reported a heritability of 22% for invasive cervix cancer and 13% for *in situ* cervix cancer.²⁴

Two studies explored cervical cancer in large twin cohorts. The Swedish Twin Registry was linked to the Swedish Cancer Registry using two age cohorts. For cervical cancer, the number of malignant cases was too small to allow for estimation of heritability. When exploring the *in situ* cases a moderate heritable effect was found. In the young cohort ($n=6900$ female pairs born between 1926 and 1958) the relative risk was 4.8 in MZ pairs and 2.4 in DZ pairs. In the older cohort ($n=5871$ female pairs born between 1886 and 1925) there were three concordant MZ pairs and no concordant DZ twin pairs.¹² In a large sample of twins ($N=8437$ MZ and $N=15\,351$ DZ) from Sweden, Denmark and Finland only one concordant MZ twin pair and three concordant DZ twin pairs with cervical cancer were found (relative risk 2.9 for MZ and 4.5 for DZ twins). The authors concluded that cancer of the cervix uteri was not influenced by genetic factors, the shared environmental influences were estimated to be 20% and the unique environmental influences 80%.¹¹ However, in both large-scale twin studies^{11,12} the number of cases is small and, therefore, the power to distinguish genetic and shared environmental factors is low.

Table 2 Polychoric correlations among female relatives (and 95% confidence intervals) for cervical smear (corrected for age at last screening and year of diagnosis)

	Polychoric correlation	95% confidence interval
Monozygotic twins pairs	0.37	0.12–0.58
First-degree relatives	0.14	–0.01–0.29

Table 3 Genetic model-fitting results for cervical smear

Model	–2LL	df	vs	χ^2	Δdf	P	AIC	a^2	c^2	e^2
1 Full ACE	2302.350	2688					–3073.650	0.37	2.6 ^{–10}	0.63
2 E model	2315.051	2690	1	12.701	2	0.002	–3064.949	—	—	1.00
3 CE model	2304.188	2689	1	1.838	1	0.175	–3073.812	—	0.23	0.77
4 AE model	2302.350	2689	1	0	1	1.000	–3075.650	0.37	—	0.63

Abbreviations: –2LL, –2 log likelihood; χ^2 , by subtracting the –2LL of two different (nested) models yields a statistic that is asymptotically distributed as χ^2 -distribution; AIC, Akaike's Information Criterion; a lower value of AIC indicates a more parsimonious model; df, degrees of freedom; Δdf , difference in degrees of freedom between two models; P, P-value; vs, tested versus model. Cervical smear results were categorized in no abnormalities, mild abnormalities and severe abnormalities. Thresholds (=prevalences) were allowed to vary with age group (<40 or ≥ 40 years) and year of diagnosis (<1996 or ≥ 1996). a^2 , additive genetic influences; c^2 , common environmental influences; e^2 , unique environmental factors.

Most studies described above suggest that the influence of genetic factors on cervical cancer is small or absent (with heritabilities ranging from 0 to 27%). For the present study an intermediate phenotype (cervix smear abnormalities) was used and a slightly higher heritability of 37% was found.

Persistent infection of oncogenic or high-risk types of HPV, such as HPV16, is considered to be a necessary, but not sufficient cause for cervical cancer. Most HPV infections are transient, and more than 90% of the infections self-cure within a few years.^{25,26} A number of factors (genetic and/or environmental) affect the liability to infection and the ability to clear this infection or alternatively develop persistent infection and subsequent CIN and cervical carcinoma. Different polymorphic human leukocyte antigen genes might be involved in the clearance and maintenance of HPV infection and a *TP53* gene allele (codon-72) might be associated with susceptibility to HPV-associated cervical carcinogenesis.^{27–29} An association between the glutathione S-transferase mu 1 null genotype and cervical cancer was reported in an US population³⁰ and a Japanese population.³¹

Both the risk of HPV infection and the risk to develop cervical cancer are associated with lifestyle factors; for example, the use of oral contraceptives, smoking, education, physical activity, high parity, risky sexual behavior, multiple sexual partners and previous exposure to other sexually transmitted diseases.^{30,32,33} Several of these factors are considered to be environmental, but it should be noted that lifestyle factors, such as smoking, physical activity and risky sexual behavior, are influenced both by genetic and environmental factors.^{34–37}

The results of the present study show familial clustering for abnormal cervix smear that is due to shared genetic factors. The largest proportion of the variation in cervical smear abnormalities is due to unique environmental factors. Individual variations in cervix smear abnormalities could arise from differences in exposure to HPV (lifestyle) and/or differences in susceptibility to HPV once exposed. Both unique environmental factors (63%) and genetic factors (37%) have a role.

These findings are important in communication and prevention of cervix smear abnormalities. A next step is to further unravel the specific genotypes and environmental factors that cause women to be at risk. Understanding the factors that influence the susceptibility to cervix smear abnormalities (and, therefore, a higher risk to develop cervix cancer) will help to develop effective prevention programs.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by the Netherlands Organization for Scientific Research (NWO-MW 904-61-193; NWO/SPI 56-464-14192 to DB; NWO 480-04-004; NWO-VENI 451-06-004 to JV).

- 1 Anttila A, Ronco G, Clifford G *et al*: Cervical cancer screening programmes and policies in 18 European countries. *Br J Cancer* 2004; **91**: 935–941.
- 2 Gustafsson L, Ponten J, Zack M, Adami HO: International incidence rates of invasive cervical cancer after introduction of cytological screening. *Cancer Causes Control* 1997; **8**: 755–763.
- 3 Yoder L, Rubin M: The epidemiology of cervical cancer and its precursors. *Oncol Nurs Forum* 1992; **19**: 485–493.
- 4 Van der Aa MA, Pukkala E, Coebergh JW, Anttila A, Siesling S: Mass screening programmes and trends in cervical cancer in Finland and the Netherlands. *Int J Cancer* 2008; **122**: 1854–1858.
- 5 Boot HJ, Wallenburg I, Melker HE *et al*: Assessing the introduction of universal human papillomavirus vaccination for preadolescent girls in The Netherlands. *Vaccine* 2007; **25**: 6245–6256.

- 6 Couto E, Hemminki K: Heritable and environmental components in cervical tumors. *Int J Cancer* 2006; **119**: 2699–2701.
- 7 Hemminki K, Chen B: Familial Risks for Cervical Tumors in Full and Half Siblings: Etiologic Apportioning. *Cancer Epidemiol Biomarkers Prev* 2006; **15**: 1413–1414.
- 8 Hemminki K, Li X, Mutanen P: Familial risks in invasive and *in situ* cervical cancer by histological type. *Eur J Cancer Prev* 2001; **10**: 83–89.
- 9 Magnusson PKE, Lichtenstein P, Gyllenstein UB: Heritability of cervical tumours. *Int J Cancer* 2000; **88**: 698–701.
- 10 Zelmanowicz AM, Schiffman M, Herrero R *et al*: Family history as a co-factor for adenocarcinoma and squamous cell carcinoma of the uterine cervix: results from two studies conducted in Costa Rica and the United States. *Int J Cancer* 2005; **116**: 599–605.
- 11 Lichtenstein P, Holm NV, Verkasalo PK *et al*: Environmental and Heritable Factors in the Causation of Cancer – Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000; **343**: 78–85.
- 12 Ahlbom A, Lichtenstein P, Malmstrom H, Feychting M, Hemminki K, Pedersen NL: Cancer in twins: genetic and nongenetic familial risk factors. *J Natl Cancer Inst* 1997; **89**: 289–293.
- 13 Boomsma DI, Vink JM, van Beijsterveldt CEM *et al*: Netherlands Twin Register: a focus on longitudinal research. *Twin Res* 2002; **5**: 401–406.
- 14 Casparie M, Tiebosch A, Burger G *et al*: Pathology databanking and biobanking in The Netherlands a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cellular Oncology* 2007; **29**: 19–24.
- 15 van Ballegoijen M, Hermens R: Cervical cancer screening in the Netherlands. *European Journal of Cancer* 2000; **36**: 2244–2246.
- 16 Rebolj M, Van Ballegoijen M, Berkers LM, Habbema D: Monitoring a national cancer prevention program: successful changes in cervical cancer screening in the Netherlands. *Int J Cancer* 2006; **120**: 806–812.
- 17 Boomsma DI, de Geus EJC, Vink JM *et al*: Netherlands Twin Register: from twins to twin families. *Twin Research and Human Genetics* 2006; **9**: 849–857.
- 18 Bulk S, van Kemenade FJ, Rozendaal L, Meijer CJLM: The Dutch CISOA-A framework for cytology reporting increases efficacy of screening upon standardisation since 1996. *J Clin Pathol* 2004; **57**: 388–393.
- 19 Falconer DS, Mackay TFC: *Threshold Characters (Chapter 18)*, in *Quantitative Genetics*. Longman Group Ltd.: Essex, 1996.
- 20 Neale MC, Boker SM, Xie G, Maes HH: *Mx: Statistical Modeling*. 5th edn. VCU Box 900126: Richmond, VA 23298, 1999.
- 21 Boomsma DI, Busjahn A, Peltonen L: Classical twin studies and beyond. *Nat Rev Genet* 2002; **3**: 872–882.
- 22 Gök M, Heideman DA, van Kemenade FJ *et al*: HPV testing on self collected cervicovaginal lavage specimens as screening method for women who do not attend cervical screening: cohort study. *British Medical Journal* 2010; **11**: c1040.
- 23 McCredie MR, Sharples KJ, Paul D *et al*: Natural history of cervical neoplasia and risk of invasive cancer in women with 3: a retrospective cohort study. *Lancet Oncol* 2008; **9**: 425–434.
- 24 Czene K, Lichtenstein P, Hemminki K: Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish family-cancer database. *Int J Cancer* 2002; **99**: 260–266.
- 25 Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD: Natural history of cervicovaginal papillomavirus infection in young women. *N Engl J Med* 1998; **338**: 423–428.
- 26 Nobbenhuis MA, Helmerhorst TJ, van den Brule AJ *et al*: Cytological regression and clearance of high-risk human papillomavirus in women with an abnormal cervical smear. *The Lancet* 2001; **358**: 1782–1783.
- 27 de Araujo Souza PS, Villa LL: Genetic susceptibility to infection with human papillomavirus and development of cervical cancer in women in Brazil. *Mutation Research* 2003; **544**: 375–383.
- 28 Koshiol J, Hildesheim A, Gonzalez P *et al*: Common Genetic Variation in TP53 and Risk of Human Papillomavirus Persistence and Progression to CIN3/Cancer Revisited. *Cancer Epidemiol Biomarkers Prev* 2009; **18**: 1631–1637.
- 29 Simões RT, Gonçalves MA, Castelli EC *et al*: HLA-G polymorphisms in women with squamous intraepithelial lesions harboring human papillomavirus. *Mod Pathol* 2009; **22**: 1075–1082.
- 30 Au WW: Life style, environmental and genetic susceptibility to cervical cancer. *Toxicology* 2004; **198**: 117–120.
- 31 Nishino K, Sekine M, Kodama S *et al*: Cigarette smoking and glutathione S-transferase M1 polymorphism associated with risk for uterine cervical cancer. *J Obstet Gynaecol Res* 2008; **34**: 994–1001.
- 32 Bosch FX, de Sanjose S: The epidemiology of human papillomavirus infection and cervical cancer. *Dis markers* 2007; **23**: 213–227.
- 33 Cotton SC, Sharp L, Masson LF *et al*: Lifestyle and socio-demographic factors associated with high-risk HPV infection in UK women. *Br J Cancer* 2007; **97**: 133–139.
- 34 Vink JM, Willemssen G, Boomsma DI: Heritability of smoking initiation and nicotine dependence. *Behav Genet* 2005; **35**: 397–406.
- 35 Stubbe JH, Boomsma DI, Vink JM *et al*: Genetic influence on exercise participation in 37,051 twin pairs from seven countries. *PLoS ONE* 2006; **20**: e22.
- 36 Verweij K, Zietsch BP, Bailey JM, Martin NG: Shared etiology of risky sexual behaviour and adolescent misconduct: genetic and environmental influences. *Genes, Brains and Behavior* 2009; **8**: 107–113.
- 37 Zietsch BP, Verweij K, Bailey AJ, Wright MJ, Martin NG: Genetic and environmental influences on risky sexual behavior and its relationship with personality. *Behav Genet* 2010; **40**: 12–21.