

# Genetic and Environmental Factors in Invasive Cervical Cancer: Design and Methods of a Classical Twin Study

Dorothy A. Machalek,<sup>1,2</sup> John D. Wark,<sup>3</sup> Sepehr N. Tabrizi,<sup>1,2,4</sup> John L. Hopper,<sup>5</sup> Minh Bui,<sup>5</sup> Gillian S. Dite,<sup>5</sup> Alyssa M. Cornall,<sup>1,2,4</sup> Marian Pitts,<sup>6</sup> Dorota Gertig,<sup>7</sup> Bircan Erbas,<sup>8</sup> Suzanne M. Garland,<sup>1,2,4</sup> on behalf of the CeCaGeEn Study Group

<sup>1</sup>Department of Microbiology and Infectious Diseases, The Royal Women's Hospital, Melbourne, Victoria, Australia

<sup>2</sup>Murdoch Childrens Research Institute, Melbourne, Victoria, Australia

<sup>3</sup>University of Melbourne Department of Medicine, and Bone and Mineral Medicine, Royal Melbourne Hospital, Melbourne, Victoria, Australia

<sup>4</sup>Department of Obstetrics and Gynaecology, University of Melbourne, Melbourne, Victoria, Australia

<sup>5</sup>Centre for Epidemiology and Biostatistics, University of Melbourne, Victoria, Australia

<sup>6</sup>Australian Research Centre in Sex, Health and Society, La Trobe University, Melbourne, Victoria, Australia

<sup>7</sup>Victorian Cervical Cytology Registry, Melbourne, Victoria, Australia

<sup>8</sup>Department of Public Health, La Trobe University, Melbourne, Victoria, Australia

**Background:** Persistent high-risk human papillomavirus (HPV) infection is a necessary prerequisite for development of cervical cancer and its precursor lesion, high-grade squamous intraepithelial lesion (HSIL). However, HPV infection is not sufficient to drive this process, and genetic and environmental factors may also play a role. **Methods/Design:** The Cervical Cancer, Genetics and Environment Twin Study was established to investigate the environmental and genetic influences on variation in susceptibility to cervical pre-cancer in 25- to 69-year-old monozygotic (MZ) and dizygotic (DZ) twins recruited through the Australian Twin Registry. Reviews of Papanicolaou (Pap) screening histories were undertaken to identify individual women with a history of an abnormal Pap test. This was followed by detection of HPV in archival Pap smears of selected twin pairs to determine HPV persistence. Selected twin pairs also completed a detailed questionnaire on socio-demographic characteristics, sexual behavior, and HPV knowledge. In future analyses, under the assumptions of the classical twin design, case-wise concordance for persistent HPV infection and HSIL will be calculated for MZ and DZ twin pairs, and twin pairs (both MZ and DZ) who are discordant for the above outcomes will be used to assess the contributions of measured environmental risk factors. **Discussion:** The study examines factors related to HPV persistence and development of HSIL among female MZ and DZ twins. The results will contribute to our understanding of the natural history of cervical HPV infection and the relative contributions of genetic and environmental factors in disease progression.

■ **Keywords:** human papillomavirus, HPV, cervical cancer, classical twin studies, co-twin control method, concordance rate, genetic factors, twinning, cohort study

Persistent infection with high-risk human papillomavirus (HR-HPV) is recognized as a necessary cause of nearly 100% of cervical cancers (Bouvard et al., 2009; De Vuyst et al., 2009). Over recent years, mounting evidence from epidemiological studies has built an understanding of the natural history of HPV infection and disease progression. Infection with HPV is extremely common in young women in their first decade of sexual activity (Castle et al., 2005). Persistent infections and cervical pre-cancer are typically established within a decade (5–10 years) for less than 10% of new infections, and subsequent invasive cancer arises over

many years, even decades, in a minority of women with pre-cancer (Castle et al., 2011; Rodriguez et al., 2010; Schiffman et al., 2002).

RECEIVED 10 December 2015; ACCEPTED 4 October 2016. First published online 5 December 2016.

ADDRESS FOR CORRESPONDENCE: Dr Dorothy Machalek, Department of Microbiology and Infectious Diseases, The Royal Women's Hospital, Melbourne VIC 3052, Australia. E-mail: [dorothy.machalek@mcri.edu.au](mailto:dorothy.machalek@mcri.edu.au)

HR-HPV infection has long been established as the primary risk factor for cervical cancer, with HPV16 having the highest carcinogenic potential (Bouvard et al., 2009). While the first infection stage of cervical cancer pathogenesis is influenced by risk factors related to sexual behavior (Syrjanen et al., 1984) among HPV-infected women, a range of environmental and behavioral risk factors – including cigarette smoking, high parity, and long-term oral contraceptive use – have been established as important co-factors in disease persistence (Jensen et al., 2013; Kapeu et al., 2009; Wang et al., 2009). Other factors suggested to play important roles in HPV susceptibility include alterations in the cervicovaginal and mucosal immune microenvironments (Kyrgiou et al., 2016; Lee et al., 2013). However, their role as modifiers of HPV natural history is poorly understood and has not been examined in longitudinal studies (Kyrgiou et al., 2016). In addition, epidemiological studies have suggested that concomitant infections with other sexually transmitted infections (STIs), in particular human immunodeficiency virus (HIV), *Chlamydia trachomatis*, and herpes simplex virus-2, may act as important co-factors for HPV persistence (Carreira et al., 2013; de Abreu et al., 2016; Samoff et al., 2005). However, history of STIs is also an indicator of sexual behavior, and concomitant infection may simply reflect recent exposure. As such, with the exception of HIV infection, consensus for the additive impact of these factors on HPV persistence is lacking.

Hereditary host factors may also play a role in driving HPV persistence and disease progression. Previous work suggests that cervical cancer may aggregate in families. Studies of the familial clustering of cervical cancer have found up to two-fold elevated risk of cervical intraepithelial neoplasia grade 2/3 (CIN 2/3) or invasive cervical cancer in women with an affected first-degree relative (Magnusson et al., 2000; Moore et al., 2012; Zelmanowicz Ade et al., 2005). More recently, large genome-wide association studies (GWAS) have identified a number of potential host genetic markers of susceptibility to persistent HPV infection leading to cancer, but results are inconsistently reported across different studies (Chen & Gyllensten, 2015; Martinez-Nava et al., 2016).

A powerful method to investigate the role of gene and familial environment on complex traits and diseases is the classical twin study (Boomsma et al., 2002; Martin et al., 1978). Because monozygotic (MZ) twins are genetically identical and dizygotic (DZ) twins share, on average, half of their genomes, they can be studied to understand the relative contributions of genetic and environmental influences on disease risk. Under the assumptions of the classical twin model, if the concordance for the outcome measure is greater for MZ pairs than for DZ pairs, then genetic factors are likely to be important. In contrast, if the concordance for MZ and DZ pairs is similar, then environmental factors shared by twins are likely to be impor-

tant (Boomsma et al., 2002; Hopper et al., 1987). Additionally, since MZ twins share the same genotypes and also tend to share lifestyle factors, the discordant identical twin model (i.e., co-twin control method) can be used to investigate the impact of environmental factors on a trait, eliminating genetic variation as a cause of discordance (van Dongen et al., 2012).

To date, familial and twin studies of cervical cancer heritability have been limited in their capacity to account for potential confounding that may explain all or some of the observed associations, for example shared cervical cancer screening practices or similar behavior patterns within families (Guo, 2001; Magnusson et al., 2000; Treloar et al., 1999), and no study has examined environmental influences on HPV persistence using a discordant-twin model. In addition, HPV infection status has not been adequately addressed by any twin study aimed at quantifying the contributions of genetic and environmental factors on risk.

The Cervical Cancer, Genetics and Environment Twin Study (CeCaGeEn) was established to determine the relative contributions of environmental and genetic factors in the development of invasive cervical cancer. Specifically, the study aims were to assess (1) whether genetic and environmental factors influence the development of persistent cervical HR-HPV infection; (2) for women with persistent HR-HPV, whether genetic and environmental factors influence the progression to cervical cancer, using HSIL as a surrogate for a cancer endpoint; and (3), among discordant twin pairs, what environmental factors influence development of persistent HR-HPV and progression to HSIL. Secondary to this, the study also aimed to assess the negative predictive value of HPV16 and HPV18 serology for past HPV infection in twins with no history of an abnormal Papanicolaou (Pap) smear result.

In future analyses, under the assumptions of the classical twin design, concordance for persistent HPV infection and HSIL will be calculated for MZ and DZ twin pairs. Data from twin pairs (both MZ and DZ) who are discordant for HPV infection and HSIL will be analyzed to assess the contribution of measured environmental risk factors. We hypothesize the following:

1. concordance rates for persistent infections will be greater for MZ twin pairs compared with DZ twin pairs in women who have been exposed to HR-HPV (Hypothesis 1);
2. in women with persistent HR-HPV, HSIL concordance rates will be greater for MZ twin pairs compared with DZ twin pairs (Hypothesis 2);
3. for discordant MZ and DZ twin pairs, the odds of persistent infection or HSIL will be higher for those with known cervical cancer risk factors such as smoking, high parity, oral contraceptive pill, and history of multiple sexual partners and STIs (Hypothesis 3).

## Design and Methods

### Design

CeCaGeEn is a retrospective study of cervical HPV infection and squamous intraepithelial neoplasia in Pap smears in 25- to 69-year-old female MZ and DZ twins recruited from the Australian Twin Registry (ATR) (Clifford & Hopper, 1986). Reviews of Pap screening histories were performed to identify individual women with a history of an abnormal Pap test. This was followed by detection of HPV in archival Pap smears of selected twin pairs to determine HPV persistence. Selected study participants also completed a detailed questionnaire capturing data on socio-demographic characteristics, sexual behavior, knowledge and understanding of HPV, and reproductive history.

### Ethics

Approval for the study was granted on December 13, 2006 by the Royal Women's Hospital (Melbourne, Australia) Human Research and Ethics Committees (Reference #11/15). The study has been conducted in accordance with the National Health and Medical Research Council's National Statement of Ethical Conduct in Human Research 2007 (<http://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research>) and the World Medical Association Declaration of Helsinki (October 2000). Written informed consent was obtained from all women before any study-specific procedures were performed.

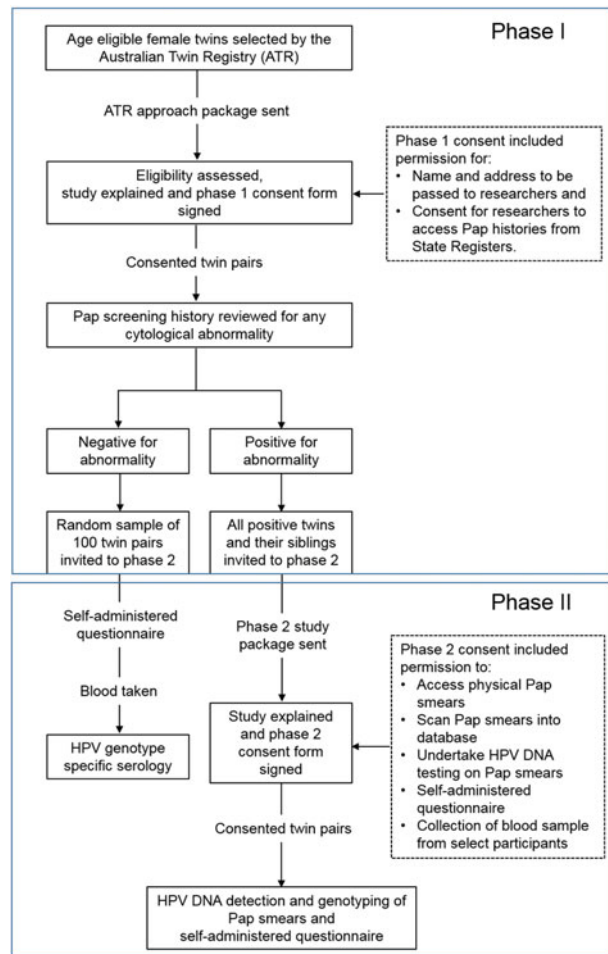
### Study Population and Recruitment

The study utilized the valuable resource of the ATR, a voluntary registry of over 35,000 twin pairs willing to take part in medical research. The ATR is funded by the Australian Government through the National Health and Medical Research Council and aims to connect researchers with twins who have volunteered to take part in medical research (Clifford & Hopper, 1986).

Female twin pairs were invited to take part in the study in a two-part process described below. The overall study flow is shown in Figure 1.

**Phase I.** Between 2007 and 2012, the ATR approached female MZ and DZ twins to participate in the study via an approach package that was mailed directly to members. Eligible twins who had not been approached to participate in research in the preceding six months were randomly selected and sent a study invitation package. Twins who did not respond to the initial approach package within two to three weeks were followed up, either with a reminder letter or a telephone call from one of the ATR's research assistants. The ATR provided contact details of the twins to the study investigators only when both twins had agreed to participate in the study and the twins were assumed to be eligible for the study.

Eligible participants were women aged from 25 to 69 years who had a twin sister. Participants were excluded if



**FIGURE 1**

(Colour online) Overview of study recruitment and procedures.

either twin was unable or declined to provide informed consent, had previously received an HPV vaccine, or were unable to complete the study requirements.

For all consenting twin pairs, Pap screening histories and pathology reports (where relevant) were obtained from cervical cytology registries. Australia has well-organized state-based registries for cervical cytology that were established in the late 1980s and 1990s under state legislation. Reporting laboratories forward all Pap smear results to the state registries, unless a woman opts out by informing her doctor or the register. Less than 1% of Pap tests are not recorded on a registry. In all states, pathology laboratories are required by the National Pathology Accreditation Advisory Council to keep Pap smear slides for at least 10 years (Smith et al., 2003a), and may keep them for longer if storage permits.

Following review of full screening histories, individual women with any cytological abnormality and their twins (regardless of Pap history) were invited to participate in Phase II of the study. In addition, 100 randomly selected twin pairs with no history of cytological abnormality were

invited to participate in Phase II. HPV16 and HPV18 serology was performed in this group to determine the negative predictive value of a negative screening history in determining past infection with HPV16 and HPV18, and in doing so evaluate how much misclassification may be present in the analysis. Involving twin pairs with normal Pap smears also allowed the study to preserve the privacy of twins who were invited to participate in Phase II (i.e., communicating to the twins that they have been selected for Phase II, but not specifically because they have had an abnormal Pap smear result).

**Phase II.** In Phase II, once informed consent was obtained, selected Pap slides were retrieved from pathology laboratories. Pathology laboratories were sent a list containing the full names, dates of birth, and unique laboratory and Medicare numbers of the participants who consented to having their slides released to the study. Slides were retrieved and delivered to the Royal Women's Hospital Department of Molecular Microbiology (Melbourne, Australia) for HPV DNA detection and genotyping (Tabrizi et al., 2010).

**Slide selection.** Following review of individual screening histories, slides with the earliest diagnosis of a cervical abnormality (atypical squamous cells of undetermined significance [ASC-US] or higher) were denoted as the index Pap and retrieved for HPV DNA analysis. In twin pairs where only one twin (Twin A) had a history of abnormality, the index Pap for her twin (Twin B) was defined as the Pap smear closest to the date of Twin A's diagnosis. Up to two Pap results directly preceding the index Pap and performed at least 12 months apart were also retrieved. HPV DNA testing was performed on the index Pap (denoted as Pap 1) and the preceding Pap slides.

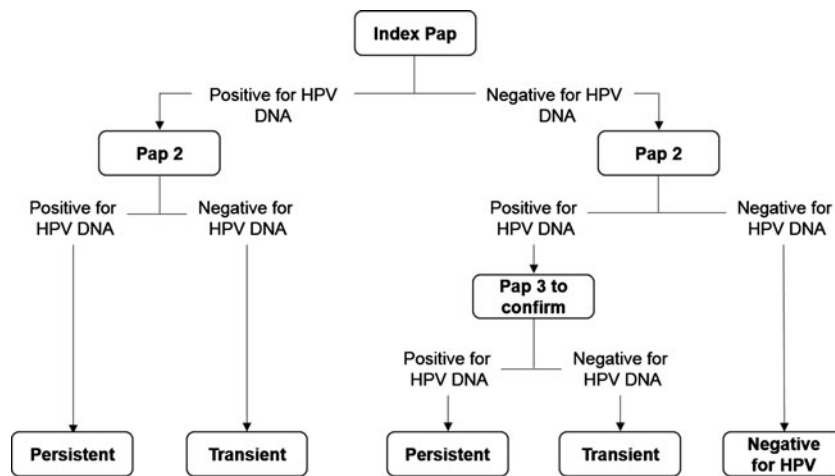
**Blood sample collection.** Blood samples were drawn from the randomly selected twin pairs with no history of cytological abnormality. These samples were used for HPV16 and HPV18 genotype serology and may be also used in the future to measure other potential determinants of cervical cancer risk.

**Questionnaire.** Phase II eligible twin pairs were also asked to complete a short paper-based questionnaire that asked about socio-demographic characteristics, smoking status, age at first sexual intercourse, number of lifetime and recent sexual partners (previous 5 years and 12 months), current contraceptive use, and relationship status. Additional information included history of *Chlamydia trachomatis* or genital herpes simplex virus type 2 infection and knowledge and understanding of HPV. A brief reproductive history was also collected (Smith et al., 2003a). The measures and scales were adopted from a number of studies, including the Australian Study of Health and Relationships (Smith et al., 2003a, 2003b) and from published HPV knowledge

and awareness surveys developed by authors MP and SMG (Pitts et al., 2007; Smith et al., 2009).

**HPV DNA detection and genotyping.** All Pap slides were first digitally scanned using an Aperio ScanScope XT scanner (Leica Biosystems, Heidelberg, Germany) before processing. Coverslips and cells were removed from the slides, and DNA was extracted on the MagNAPure LC automated system using the DNA-I kit (blood cells high-performance protocol), as previously described (Chua & Hjerpe, 1996; Tabrizi et al., 2010). DNA samples were screened for the presence of HPV DNA using the DNA ELISA kit HPV SPF10, version 1 (Labo Bio-medical Products BV, Rijswijk, The Netherlands), following the manufacturer's instructions. Samples that were negative for HPV DNA were screened for a 110 bp segment of the human beta-globin gene by quantitative real-time PCR, as previously described (Cornall et al., 2013). Samples that were negative for both HPV and beta-globin DNA were considered not assessable and omitted from further analysis. HPV DNA positive samples were amplified, detected, and scored on the LINEAR ARRAY<sup>®</sup> HPV Genotyping Test (Roche Diagnostics GmbH, Mannheim, Germany) following the manufacturer's instructions and with previously described modifications (Stevens et al., 2008). DNA samples that were invalid on the LINEAR ARRAY<sup>®</sup> HPV Genotyping Test (negative HPV and for one or both beta-globin probes) were tested on the RHA kit HPV SPF10-LiPA25, version 1 (Labo Bio-medical Products BV, Rijswijk, The Netherlands) following the manufacturer's instructions and using the same amplicons generated for the DNA ELISA. Positive results for HPV34, HPV43, HPV44, and HPV74 on HPV SPF10-LiPA25 were excluded from analysis because these genotypes are unable to be detected by LINEAR ARRAY<sup>®</sup> HPV Genotyping Test. Because HPV SPF10-LiPA25 is unable to discriminate between HPV68 and HPV73 (Kleter et al., 1999), samples with this result were further tested on type-specific quantitative real-time PCR for HPV68 and HPV73, as previously described (Cornall et al., 2015). All other HPV SPF10-LiPA25 positive results were recorded.

**HPV pseudovirion (PsV) neutralization assays.** Blood samples were transported to the Diamantina Institute (University of Queensland, Australia) for processing and testing for the presence of HPV16 and HPV18 antibodies using HPV PsV neutralization assays. PsV16 and PsV18, encapsidating a secreting alkaline phosphatase reporting plasmid produced in 293TT cells, were used for in vitro neutralization assays using well-established methods, as previously described (Buck et al., 2005; Pastrana et al., 2015). Infection of 293TT cells was monitored by secreting alkaline phosphatase activity in the culture supernatant using a highly sensitive chemiluminescent reporter system. Antibody-mediated PsV neutralization was detected by a reduction in secreting alkaline phosphatase activity. The titre was defined as the reciprocal of the highest dilution of



**FIGURE 2**  
Definition of participant status based on HPV DNA testing of Pap slides.

serum that reduced the secreting alkaline phosphatase activity by at least 50% compared with the reactivity in the wells that received PsV but no antibody (Pastrana et al., 2015).

**Statistical considerations. Outcome definition.** Following HPV testing, Pap screening histories of individual members of a twin pair were classified as persistent, transient, or negative for HPV (Figure 2). Persistent HPV infection was defined as detection of the sample HPV genotype in two Pap smear samples collected over at least a 12-month period, regardless of the cytology result recorded for the index Pap smear (e.g., normal/LSIL/HSIL). Transient infection was defined as having only one Pap slide positive for HPV DNA with Normal/LSIL abnormality recorded on the index Pap smear. Serial Pap slides with no history of cytological abnormality and on which HPV DNA was not detected were classified as negative for HPV (Figure 2).

High-grade lesions were defined as ASC-cannot exclude HSIL (ASC-H) or HSIL recorded on the Pap test Registry plus HSIL confirmed by biopsy.

**Analysis plan.** The main objective of the analyses is to conduct case-wise concordance analyses to examine whether the development of persistent infection with HR-HPV (Aim 1) or further progression to cervical cancer using cytological abnormalities or HSIL as a surrogate (Aim 2) is influenced by genetic or environmental factors, after adjusting for other known important covariates. Case-wise concordance is the conditional probability of a twin being affected, given that the co-twin is also affected (Witte et al., 1999). Under the assumptions of the classical twin model, non-genetic factors relevant to the trait variance are shared to the same extent within MZ pairs as within DZ pairs. If there is a familial influence on risk, the concordance within

twin pairs will be greater than expected by chance; if concordance for a trait is greater within MZ pairs than within DZ pairs, the presence of an additive genetic influence is supported; if concordance within pairs is not different between MZ and DZ pairs, then a shared environmental influence is suggested. To address Aim 1, case-wise concordance for persistent versus transient infection (Figure 2) will be calculated among MZ and DZ pairs. For Aim 2, the analysis will be restricted to MZ and DZ twins with persistent HR-HPV infection (Figure 2). Case-wise concordance will be estimated using maximum likelihood methods with 95% confidence intervals, whereas comparisons between MZ and DZ pairs will be performed using chi-square statistics. Both univariate and adjusted analysis will be performed, adjusting for known cervical cancer risk factors such as HPV status, smoking, high parity, oral contraceptive pill, and history of multiple sexual partners and STIs.

MZ and DZ pairs who are discordant for persistent infection or HSIL will be used to assess the effects of environmental risk factors for cervical cancer pathogenesis. Logistic regression estimated by generalized estimating equations (which take into account the correlation within twin pairs) will be used to examine association between cervical cancer pathogenesis and independent variables, including zygosity, smoking, high parity, oral contraceptive use, history of sexual transmitted infections (e.g., *Chlamydia trachomatis* and genital herpes), and plus any other potential confounders.

All statistical tests will be two sided with significance level chosen at 5%. All analyses will be performed using STATA version 13 (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP) and an in-house program written in R (<https://www.r-project.org/>).

**Power and sample size estimates.** Given the limited data available from previous studies, it is difficult to predict an effect size on which to base the sample size (Moore et al., 2012). Nevertheless, we expect a small difference in concordance between MZ and DZ pairs. To determine sample sizes, tests of two sample independent proportions were performed. The analysis showed that a total of 72 twin pairs is required in each of the groups (assuming equal numbers of MZ and DZ pairs) to detect a difference in concordance of 0.18 on the primary outcomes, given a two-sided 5% significance level, a power of 80%, and a case-wise concordance of 0.3 for MZ twin pairs. However, under the assumptions that 50% of invited twin pairs consent, 50% have Pap smears available, an HR-HPV prevalence of 25%, and a 0.1% probability that one or both twins have a history of HSIL, a substantially larger number of twin pairs are needed ( $n = 364$  to address Aim 1 and  $n = 198$  to address Aim 2).

## Discussion

The CeCaGeEn study is a multidisciplinary collaboration bringing together a wealth of experience in the fields of molecular biology, epidemiology, behavioral science, and twin research. The study uses the resources of the ATR and the cervical cytology registries to examine the environmental and genetic influences on variation in susceptibility to cervical pre-cancer. In addition, effects of environmental risk factors for cervical cancer pathogenesis will be investigated among discordant twin pairs. The study results will contribute to understanding the natural history of cervical HPV infection and the relative contribution of genetic and environmental factors in disease progression.

The determinants of HR-HPV persistence and subsequent progression from HSIL to cervical cancer are unknown. Only a small proportion of HPV-infected women ever develop cervical cancer, and a combination of viral, environmental, and host genetic factors are likely to be important determinants of disease risk. With respect to the latter, the contribution of heritable factors to the causation of cervical cancer has previously been quantified, with estimates ranging from 22% to 27% based on two large studies of familial cancer aggregation (Czene et al., 2002; Magnusson et al., 2000), and a recent GWAS (Chen et al., 2015). These estimates of contribution of genetic factors for cervical cancer were substantially higher than that for colon (13%), lung (8%), and kidney (7%) cancers (Czene et al., 2002).

A large proportion of susceptibility to cervical cancer pathogenesis (>70%) may not be attributed to additive genetic factors; therefore, a better understanding of other contributing factors is critical. The importance of such risk factors for HPV persistence can be investigated using the co-twin control design, using twins discordant for a specific environmental factor (i.e., history of chlamydia infection, smoking, and oral contraceptive use), as well as twins discordant for disease outcome (i.e., HR-HPV infection or per-

sistence). The co-twin control method is an elegant research design that uses differences within twin pairs to examine the association between a measured risk factor and an outcome of interest, which improves the power of the conventional unmatched (case-control study) design (Goldberg & Fischer, 2005). The advantage of using the co-twin control method is that it provides a powerful tool to control for genetic and unmeasured early environmental factors. The results of such analyses may provide novel data relating to the etiology of cervical cancer.

This study is subject to potential limitations. First, HPV serology was limited to HPV genotypes 16 and 18; therefore, we cannot assess previous exposure to other HR-HPV genotypes. However, HPV16 and 18 are responsible for the majority of cervical disease and are perhaps the most relevant in the context of assessing pre-vaccination seroprevalence rates among women with normal screening histories (Bouvard et al., 2009; Velentzis et al., 2014). Also, only a small number of blood samples were collected, which limits our power to perform additional risk-factor or twin analyses. Second, recent studies indicate that the cervicovaginal microbiome may play an important role in susceptibility to STI infections (including HPV) and to the development of cervical HSIL. While we did not collect vaginal microbiome samples, the DNA samples extracted for this study will likely be used in future sub-studies, including cervical microbiome profiling (Mittra et al., 2015; Smith et al., 2012). Third, HPV detection and genotyping in archival cervical smears can potentially yield low-quality and low-quantity DNA, depending on factors such as smear cellularity, age, and storage history (Griffin & Doorbar, 2016). Despite this, previous research suggests that HPV testing in archival cervical smears can offer accuracy in HPV detection, when compared to traditionally collected samples (Tabrizi et al., 2010). Last, we acknowledge that the contribution of additive genetic factors to risk of persistence could be overestimated due to the presence of unknown confounders.

However, the study has a number of strengths, including the availability of data on lifetime cervical screening history, as well as demographic and risk factor information and samples for HPV genotypes. Very few classical twin studies have assessed the evidence of a possible genetic component in cervical cancer pathogenesis, and none has incorporated HPV status and other potential behavioral confounders into the study design (Moore et al., 2012). As such, results of studies have been mostly inconsistent, owing to differences in study design, assessment of exposure variables, and analyses.

Identification of predictive biomarkers (whether genetic or environmental) that could determine individuals' susceptibility to cervical cancer has important clinical implications. HR-HPV testing is considered superior to Pap testing for detecting underlying disease (Castle et al., 2012; Ronco et al., 2014). As a result, the use of routine Pap testing is being phased out and replaced with HR-HPV primary

screening in many countries with organized screening programs (Castle et al., 2012). In addition, persistent HR-HPV infection has been identified as an appropriate endpoint for efficacy studies of new-generation prophylactic HPV vaccines (Lowy et al., 2015; Syrjanen, 2011). Given the evolution in clinical practice, a better understanding of all the factors that influence the risk of HPV persistence is needed. This information could be used to better define primary endpoints in novel prophylactic vaccine trials, and in clinical screening algorithms to pinpoint those most at risk and requiring further follow-up.

## Acknowledgments

Funding for this study was provided by the National Health and Medical Research Council project grant (509002) and Cancer Australia grant (1008811). The CeCaGeEn Study Group includes Suzanne M. Garland, John D Wark, Sepehr Tabrizi, John L. Hopper, Dorota Gertig, Marian Pitts, Bircan Erbas, Ian Frazer, Michael McCullough, Dorothy A. Machalek, Minh Bui, Alyssa Cornall, Samuel Phillips, and Gillian S. Dite. We sincerely thank the following research assistants and project managers over the years; Gemma Christie, Elya Moore, Nicole Taylor, Amanda Richardson, Nicole Lister, Juliana Tasevska, Anna Debenham, Melissa Sullivan, Debra Tamvakis, Melissa Yow, and Houda Abdo. Also, we express our thanks to staff of the Australian Twin Registry and Victorian Cytology Registry and other pathology laboratories for retrieving slides. Finally, we thank all the participants without whom the study could not have happened.

## Authors' Contributions

SMG and JDW conceived of the study. SMG, JDW, JLH, SNT, DG, MP, and BE were involved in the design of the study. DAM drafted the manuscript. GSD, MB, and AMC helped draft the manuscript. All authors reviewed the manuscript for important intellectual content.

## Disclosure of Interests

SMG has received grants to her institution from Commonwealth Department of Health for HPV genoprevalence surveillance post vaccination, Merck and GSK to perform phase 3 clinical vaccine trials, Merck to evaluate HPV in RRP post vaccination program, CSL for HPV in cervical cancer study, and VCA for a study on effectiveness of public health HPV vaccine study plus a study on associations of early onset cancers. She also received speaking fees from MSD and SPMSD for work performed in her personal time. Merck paid for travel and accommodation to present at HPV Advisory board meetings. DG is the principal investigator of the NHMCR-funded iPap trial and a co-investigator on the Compass trial. The Compass trial received a funding contribution from Roche Molecular Sys-

tems USA. All the other authors declare that they have no competing interests.

## References

- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, 3, 872–882.
- Bouvard, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F., ... Coglianò, V. (2009). A review of human carcinogens — Part B: Biological agents. *Lancet Oncology*, 10, 321–322.
- Buck, C. B., Pastrana, D. V., Lowy, D. R., & Schiller, J. T. (2005). Generation of HPV pseudovirions using transfection and their use in neutralization assays. *Methods in Molecular Medicine*, 119, 445–462.
- Carreira, H., Coutinho, F., Carrilho, C., & Lunet, N. (2013). HIV and HPV infections and ocular surface squamous neoplasia: Systematic review and meta-analysis. *British Journal of Cancer*, 109, 1981–1988.
- Castle, P. E., de Sanjose, S., Qiao, Y. L., Belinson, J. L., Lazcano-Ponce, E., & Kinney, W. (2012). Introduction of human papillomavirus DNA screening in the world: 15 years of experience. *Vaccine*, 30(Suppl. 5), F117–122.
- Castle, P. E., Rodriguez, A. C., Burk, R. D., Herrero, R., Wacholder, S., Hildesheim, A., ... Proyecto Epidemiologico Guanacaste Group. (2011). Long-term persistence of prevalently detected human papillomavirus infections in the absence of detectable cervical precancer and cancer. *Journal of Infectious Diseases*, 203, 814–822.
- Castle, P. E., Schiffman, M., Herrero, R., Hildesheim, A., Rodriguez, A. C., Bratti, M. C., ... Burk, R. D. (2005). A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste, Costa Rica. *Journal of Infectious Diseases*, 191, 1808–1816.
- Chen, D., Cui, T., Ek, W. E., Liu, H., Wang, H., & Gyllensten, U. (2015). Analysis of the genetic architecture of susceptibility to cervical cancer indicates that common SNPs explain a large proportion of the heritability. *Carcinogenesis*, 36, 992–998.
- Chen, D., & Gyllensten, U. (2015). Lessons and implications from association studies and post-GWAS analyses of cervical cancer. *Trends in Genetics*, 31, 41–54.
- Chua, K. L., & Hjerpe, A. (1996). Persistence of human papillomavirus (HPV) infections preceding cervical carcinoma. *Cancer*, 77, 121–127.
- Clifford, C. A., & Hopper, J. L. (1986). The Australian NHMRC Twin Registry. A resource for the Australian scientific community. *Medical Journal of Australia*, 145, 63–65.
- Cornall, A. M., Roberts, J. M., Garland, S. M., Hillman, R. J., Grulich, A. E., & Tabrizi, S. N. (2013). Anal and perianal squamous carcinomas and high-grade intraepithelial lesions exclusively associated with 'low-risk' HPV genotypes 6 and 11. *International Journal of Cancer*, 133, 2253–2258.
- Cornall, A. M., Roberts, J. M., Molano, M., Machalek, D. A., Phillips, S., Hillman, R. J., ... Team, S. S. (2015). Laser capture microdissection as a tool to evaluate human papillomavirus genotyping and methylation as biomarkers of

- persistence and progression of anal lesions. *BMJ Open*, 5, e008439.
- Czene, K., Lichtenstein, P., & Hemminki, K. (2002). Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish family-cancer database. *International Journal of Cancer*, 99, 260–266.
- de Abreu, A. L., Malaguti, N., Souza, R. P., Uchimura, N. S., Ferreira, E. C., Pereira, M. W., ... Consolaro, M. E. (2016). Association of human papillomavirus, Neisseria gonorrhoeae and Chlamydia trachomatis co-infections on the risk of high-grade squamous intraepithelial cervical lesion. *American Journal of Cancer Research*, 6, 1371–1383.
- De Vuyst, H., Clifford, G. M., Nascimento, M. C., Madeleine, M. M., & Franceschi, S. (2009). Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: A meta-analysis. *International Journal of Cancer*, 124, 1626–1636.
- Goldberg, J., & Fischer, M. (2005). Co-twin control methods. In Brian S. Everitt & David C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*. New York: John Wiley & Sons <http://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa143/full>.
- Griffin, H., & Doorbar, J. (2016). Detection of papillomavirus gene expression patterns in tissue sections. *Current Protocols in Microbiology*, 41, 14B17.11–14B17.20.
- Guo, S. W. (2001). Does higher concordance in monozygotic twins than in dizygotic twins suggest a genetic component?. *Human Heredity*, 51, 121–132.
- Hopper, J. L., Derrick, P. L., & Clifford, C. A. (1987). Innovations in the statistical analysis of twin studies. *Acta Geneticae Medicae et Gemellologiae (Roma)*, 36, 21–27.
- Jensen, K. E., Schmiedel, S., Norrild, B., Frederiksen, K., Iftner, T., & Kjaer, S. K. (2013). Parity as a cofactor for high-grade cervical disease among women with persistent human papillomavirus infection: A 13-year follow-up. *British Journal of Cancer*, 108, 234–239.
- Kapeu, A. S., Luostarinen, T., Jellum, E., Dillner, J., Hakama, M., Koskela, P., ... Lehtinen, M. (2009). Is smoking an independent risk factor for invasive cervical cancer? A nested case-control study within Nordic biobanks. *American Journal of Epidemiology*, 169, 480–488.
- Kleter, B., van Doorn, L. J., Schrauwen, L., Molijn, A., Sastrowijoto, S., ter Schegget, J., ... Quint, W. (1999). Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. *Journal of Clinical Microbiology*, 37, 2508–2517.
- Kyrgiou, M., Mitra, A., & Moscicki, A. B. (2016). Does the vaginal microbiota plays a role in the development of cervical cancer? *Translational Research*, S1931–5244, 30109–30118.
- Lee, J. E., Lee, S., Lee, H., Song, Y. M., Lee, K., Han, M. J., ... Ko, G. (2013). Association of the vaginal microbiota with human papillomavirus infection in a Korean twin cohort. *PLoS One*, 8, e63514.
- Lowy, D. R., Herrero, R., Hildesheim, A., for the Participants in the IARC/NCI Workshop on Primary Endpoints for Prophylactic HPV Vaccine Trials. (2015). Primary endpoints for future prophylactic human papillomavirus vaccine trials: Towards infection and immunobridging. *Lancet Oncology*, 16, e226–233.
- Magnusson, P. K., Lichtenstein, P., & Gyllensten, U. B. (2000). Heritability of cervical tumours. *International Journal of Cancer*, 88, 698–701.
- Martin, N. G., Eaves, L. J., Kearsley, M. J., & Davies, P. (1978). The power of the classical twin study. *Heredity (Edinburgh)*, 40, 97–116.
- Martinez-Nava, G. A., Fernandez-Nino, J. A., Madrid-Marina, V., & Torres-Poveda, K. (2016). Cervical cancer genetic susceptibility: A systematic review and meta-analyses of recent evidence. *PLoS One*, 11, e0157344.
- Mitra, A., MacIntyre, D. A., Lee, Y. S., Smith, A., Marchesi, J. R., Lehne, B., ... Kyrgiou, M. (2015). Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. *Scientific Reports*, 5, 16865.
- Moore, E. E., Wark, J. D., Hopper, J. L., Erbas, B., Garland, S. M., & CeCaGeEn Study, G. (2012). The roles of genetic and environmental factors on risk of cervical cancer: A review of classical twin studies. *Twin Research and Human Genetics*, 15, 79–86.
- Pastrana, D. V., Buck, C. B., Lowy, D. R., & Schiller, J. T. (n.d.). Papillomavirus neutralization assay. Retrieved October 7, 2015, from <http://home.ccr.cancer.gov/LCO/neutralizationassay.htm>
- Pitts, M. K., Dyson, S. J., Rosenthal, D. A., & Garland, S. M. (2007). Knowledge and awareness of human papillomavirus (HPV): Attitudes towards HPV vaccination among a representative sample of women in Victoria, Australia. *Sex Health*, 4, 177–180.
- Rodriguez, A. C., Schiffman, M., Herrero, R., Hildesheim, A., Bratti, C., Sherman, M. E., ... Burk, R. D. (2010). Longitudinal study of human papillomavirus persistence and cervical intraepithelial neoplasia grade 2/3: Critical role of duration of infection. *Journal of the National Cancer Institute*, 102, 315–324.
- Ronco, G., Dillner, J., Elfstrom, K. M., Tunesi, S., Snijders, P. J., Arbyn, M., ... Meijer, M. D. for the International HPV Screening Working Group. (2014). Efficacy of HPV-based screening for prevention of invasive cervical cancer: Follow-up of four European randomised controlled trials. *Lancet*, 383, 524–532.
- Samoff, E., Koumans, E. H., Markowitz, L. E., Sternberg, M., Sawyer, M. K., Swan, D., ... Unger, E. R. (2005). Association of Chlamydia trachomatis with persistence of high-risk types of human papillomavirus in a cohort of female adolescents. *American Journal of Epidemiology*, 162, 668–675.
- Schiffman, M., Wheeler, C. M., Castle, P. E., & A typical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesion Triage Study Group. (2002). Human papillomavirus DNA remains detectable longer than related cervical cytologic abnormalities. *Journal of Infectious Diseases*, 186, 1169–1172.



- Smith, A. M., Rissel, C. E., Richters, J., Grulich, A. E., & de Visser, R. O. (2003a). Sex in Australia: Reproductive experiences and reproductive health among a representative sample of women. *Australian and New Zealand Journal of Public Health*, *27*, 204–209.
- Smith, A. M., Rissel, C. E., Richters, J., Grulich, A. E., & de Visser, R. O. (2003b). Sex in Australia: The rationale and methods of the Australian study of health and relationships. *Australian and New Zealand Journal of Public Health*, *27*, 106–117.
- Smith, A., Lyons, A., Pitts, M., Croy, S., Ryall, R., Garland, S., ... Tay, E. H. (2009). Assessing knowledge of human papillomavirus and collecting data on sexual behavior: Computer assisted telephone versus face to face interviews. *BMC Public Health*, *9*, 429.
- Smith, B. C., McAndrew, T., Chen, Z., Harari, A., Barris, D. M., Viswanathan, S., ... Burk, R. D. (2012). The cervical microbiome over 7 years and a comparison of methodologies for its characterization. *PLoS One*, *7*, e40425.
- Stevens, M. P., Garland, S. M., & Tabrizi, S. N. (2008). Development and validation of a real-time PCR assay specifically detecting human papillomavirus 52 using the Roche Light-Cycler 480 system. *Journal of Virological Methods*, *147*, 290–296.
- Syrjanen, K. (2011). Persistent high-risk human papillomavirus (HPV) infections as surrogate endpoints of progressive cervical disease. Potential new endpoint for efficacy studies with new-generation (non-HPV 16/18) prophylactic HPV vaccines. *European Journal of Gynaecological Oncology*, *32*, 17–33.
- Syrjanen, K., Vayrynen, M., Castren, O., Yliskoski, M., Mantyjarvi, R., Pyrhonen, S., & Saarikoski, S. (1984). Sexual behaviour of women with human papillomavirus (HPV) lesions of the uterine cervix. *British Journal of Venereal Diseases*, *60*, 243–248.
- Tabrizi, S. N., Taylor, N., McCullough, M. J., Phillips, G., Wark, J., Gertig, D., ... CeCaGeEn Study, G. (2010). Human papillomavirus genotype detection from archival papanicolaou-stained cervical tests. *Cancer Cytopathology*, *118*, 482–489.
- Treloar, S. A., McDonald, C. A., & Martin, N. G. (1999). Genetics of early cancer detection behaviours in Australian female twins. *Twin Research*, *2*, 33–42.
- van Dongen, J., Slagboom, P. E., Draisma, H. H., Martin, N. G., & Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. *Nature Reviews Genetics*, *13*, 640–653.
- Velentzis, L. S., Sitas, F., O'Connell, D. L., Darlington-Brown, J., Egger, S., Sinha, R., ... Canfell, K. (2014). Human papillomavirus 16/18 seroprevalence in unvaccinated women over 30 years with normal cytology and with high grade cervical abnormalities in Australia: Results from an observational study. *BMC Infectious Diseases*, *14*, 3861.
- Wang, S. S., Zuna, R. E., Wentzensen, N., Dunn, S. T., Sherman, M. E., Gold, M. A., ... Walker, J. L. (2009). Human papillomavirus cofactors by disease progression and human papillomavirus types in the study to understand cervical cancer early endpoints and determinants. *Cancer Epidemiology, Biomarkers & Prevention*, *18*, 113–120.
- Witte, J. S., Carlin, J. B., & Hopper, J. L. (1999). Likelihood-based approach to estimating twin concordance for dichotomous traits. *Genetic Epidemiology*, *16*, 290–304.
- Zelmanowicz Ade, M., Schiffman, M., Herrero, R., Goldstein, A. M., Sherman, M. E., Burk, R. D., ... Hildesheim, A. (2005). 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. *International Journal of Cancer*, *116*, 599–605.