**EPIDEMIOLOGIC CHARACTERIZATION OF HUMAN POLYOMAVIRUSES IN ADULTS FROM THE UNITED STATES, AND THEIR RELATION TO CUTANEOUS SQUAMOUS CELL CARCINOMA**

A Thesis

Submitted to the Faculty

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

in

Quantitative Biomedical Sciences

by

Anala Gossai

DARTMOUTH COLLEGE

Hanover, New Hampshire

March 2016

Examining Committee:

(chair) ***Margaret R. Karagas, Ph.D.***

***Brock C. Christensen, Ph.D.***

***Anne G. Hoen, Ph.D.***

***Heather H. Nelson, Ph.D.***

F. Jon Kull, Ph.D.
Dean of Graduate Studies

# Abstract

Polyomaviruses (PyV) are potentially tumorigenic viruses in humans. However, limited data exists on the population seroprevalence or longitudinal serostability of PyVs, and individual characteristics that relate to seropositivity. Further, PyVs may be associated with the occurrence of cutaneous squamous cell carcinoma (SCC).

We determined the seroprevalence of ten human PyVs among participants from a US population-based case-control study. Antibody response against each PyV type was measured using multiplex serology of recombinantly expressed VP1 capsid proteins in 253 SCC cases and 460 matched controls. On a subset of control participants, lymphocyte subtype proportions were inferred from DNA methylation profiles. Amongst controls, JC and TSV seropositivity was more common among men than women; smokers were more likely to be HPyV9 seropositive but MCV seronegative; and HPyV7 seropositivity was associated with prolonged glucocorticoid use. Differences were observed in CD8+ T and B cell proportions by BK, JC, and HPyV9 serostatus. Odds ratios (OR) for SCC incidence associated with seropositivity to each PyV type were estimated using unconditional logistic regression. Those who were JC seropositive had increased odds of SCC when compared to JC seronegativity (OR=1.37, 95% CI: 0.98-1.90), with an increasing trend in risk with increasing quartiles of seroreactivity (*P*-for-trend=0.04).

In a US nested case-control study, BK and JC seroreactivity was measured on 113 SCC cases and 229 matched controls who had a prior keratinocyte cancer. Repeated serum samples from controls, and both pre- and post-diagnosis samples from a subset of SCC cases, were also assayed. Among controls, BK and JC seroreactivity was stable over time, and there was little evidence of seroconversion following SCC diagnosis among cases. Odds of SCC associated with

seropositivity to each PyV type were estimated using conditional logistic regression. JC seropositivity prior to diagnosis was associated with an elevated risk of SCC (OR=2.54, 95% CI: 1.23-5.25).

Our findings suggest PyV seropositivity is common in the US, and varies by sociodemographic and immunologic characteristics. PyV antibody levels were stable over time and following an SCC diagnosis. This thesis also provides some evidence that a history of PyV infection may be related to the occurrence of SCC in the US population.

# Acknowledgements

Firstly, I would like to thank my advisor, Dr. Margaret R. Karagas, without whose mentorship these projects would have been impossible. Secondly, I would like to thank the members of my dissertation committee and co-authors on all projects: Dr. Brock C. Christensen, Dr. Anne G. Hoen, and Dr. Heather H. Nelson. I also thank the members of my qualifying exam committee: Dr. Jennifer A. Doherty, Dr. Jiang Gui, and Dr. Richard I. Enelow for reviewing my work and offering constructive insights and suggestions. Thirdly, I'd like to extend my thanks to the faculty, staff, and students of Dartmouth's Quantitative Biomedical Sciences program. Finally, I would like to thank the study staff and participants of the New Hampshire Skin Cancer Study and the Skin Cancer Prevention Study, without whom this work would not be possible. This research was funded in part by grants NCI R01CA057494, R01CA121147 and R01CA082354; NIHGMS P20GM104416, and NIH/NLM K01LM011985.

# Table of Contents

# List of Tables

# List of Illustrations

# Abbreviations

| | |
|---|---|
| 27k array | Illumina Infinium HumanMethylation27 BeadChip array |
| AAPC | average annual percentage change |
| AK | actinic keratosis |
| BCC | basal cell carcinoma |
| BK | BK virus |
| CI | confidence interval |
| ELISA | enzyme-linked immunosorbent assay |
| EV | epidermodysplasia verruciformis |
| GAM | generalized additive model |
| GO | gene ontology |
| GST | glutathione S-transferase |
| HLA | human leukocyte antigen |
| HPV | human papillomavirus |
| HPyV6 | human polyomavirus 6 |
| HPyV7 | human polyomavirus 7 |
| HPyV9 | human polyomavirus 9 |
| HPyV10 | Malawi/human polyomavirus 10 |
| ICC | intraclass correlation coefficient |
| Ig | immunoglobulin |
| IQR | interquartile range |
| JC | John Cunningham virus |
| κ (kappa) | kappa statistic |
| KC | keratinocyte cancer |
| KI | Karolinska Institute polyomavirus |
| KIR | killer immunoglobulin-like receptors |
| LTAg | large T antigen |
| MCC | Merkel cell carcinoma |
| MCV | Merkel cell polyomavirus |
| MDR | multifactor dimensionality reduction |
| MFI | median fluorescence intensity |
| NK cells | natural killer cells |
| NKG2D | NK group 2 member D |
| NMSC | nonmelanoma skin cancer |
| NSAIDS | nonsteroidal anti-inflammatory drugs |
| OR | odds ratio |
| PAH | polycyclic aromatic hydrocarbon |
| PML | progressive multifocal leukoencephalopathy |
| PUFA | polyunsaturated fatty acid |
| PUVA | psoralen UVA |
| PVAN | polyomavirus-associated nephropathy |
| PyV | polyomavirus |
| ρ (rho) | Spearman rank correlation coefficient |
| Rb | retinoblastoma |
| RCA | rolling circle amplification |
| SCC | cutaneous squamous cell carcinoma |
| SD | standard deviation |
| SES | socioeconomic status |
| SNP | single nucleotide polymorphism |
| STAg | small T antigen |
| SV40 | simian virus 40 |
| TAg | T antigen |
| TS | *Trichodysplasia spinulosa* |
| TSV | *Trichodysplasia spinulosa*-associated polyomavirus |
| USA | United States of America |
| UVR | ultraviolet radiation |
| WHIM | warts, hypogammaglobulinemia, infections, and myelokathexis |
| WU | Washington University polyomavirus |
| XP | xeroderma pigmentosum |

# Chapter 1

# General introduction

*Adapted from: Anala Gossai, Dorothea T. Barton, Judy R. Rees, Heather H. Nelson, and Margaret R. Karagas. "Chapter 58: Keratinocyte cancers." Schottenfeld & Fraumeni's Cancer Epidemiology and Prevention, 4th ed. Oxford (2016). (in press; excluding sections 1.1 & 1.3).*

## 1.1 HUMAN POLYOMAVIRUSES

Polyomaviruses (PyV) have been discovered in humans, nonhuman primates, mice, birds, bats, and a host of other species.[1] However, the diversity of the human polyomavirome and its impact on human health has not been fully elucidated, as evidenced by the continued discovery of new PyVs within well and ill individuals. The PyV is a non-enveloped virus with an icosahedral capsid ~45 nm in diameter, containing a circular double-stranded DNA genome of ~5,000 base pairs.[2, 3] The genome of the *Polyomaviridae* family encodes three capsid proteins (VP1, VP2, VP3), as well as a small and large T antigen (TAg).[1, 2] The viral life cycle is divided into an early phase taking place before viral DNA replication, and a late phase occurring after replication of the viral genome.[4] Early in infection, the early proteins large TAg (LTAg) and small TAg (STAg) are expressed. The LTAg contains a J domain which is required for efficient viral replication, and is involved in the regulation of early and late gene expression.[4] The capsid proteins are expressed during the late phase of infection.[4] Primary human PyV infection typically occurs during childhood;[2, 5] but as yet, the potential risk factors for infection are not well understood.

Although prevalent in various tissues across multiple populations, most PyV infections have yet to be associated with a specific disease phenotype.[6] In the immunocompetent host, the virus is believed to remain latent following primary infection. In immunocompromised patients, BK virus is associated with polyomavirus-associated nephropathy (PVAN) and cystitis,[7, 8] John Cunningham virus (JC) with progressive multifocal leukoencephalopathy (PML),[7, 8] and *Trichodysplasia spinulosa*-associated polyomavirus (TSV) with the rare skin disease *Trichodysplasia spinulosa* (TS)[9] characterized by spiny lesions due to abnormal maturation of hair follicles.[10, 11] Merkel cell polyomavirus (MCV) containing mutations in LTAg [12] is causally related to the aggressive skin cancer, Merkel cell carcinomas (MCC).[12, 13, 14]

PyVs cause tumors in animal experiments and *in vitro* assays. The LTAg has transforming potential due to its ability to bind and thereby inactivate tumor suppressor proteins retinoblastoma (Rb)[15] and p53,[16] and stimulate entry into the host cell and cell cycle.[2] The STAg and agnoprotein[17] expressed by some viruses may also exhibit oncogenic properties.[3] However, there is limited support for carcinogenesis in humans.

### 1.1.1 Historical aspects and discovery

PyVs were first described in the peer-reviewed literature in the year 1953 within mice,[18] and were named for their notable ability to induce a wide range of tumors in murine models.[19] BK and JC were the first human PyVs isolated,[20, 21] and were observed to be highly homologous in amino acids. Until the year 2006, BK virus, JC virus, and simian virus 40 (SV40) – a natural infection of some species of Asiatic macaques, were the only PyVs known to circulate in the human population, which was mainly confirmed by polymerase chain reaction (PCR) and the presence of virus-specific antibodies.

Advanced high-throughput sequencing and improved rolling circle amplification (RCA) techniques rapidly identified novel human PyVs. In the year 2007, two new PyVs found within nasopharyngeal samples were described: the Karolinska Institute polyomavirus (KI) was identified by a research group at the Karolinska Institute, Sweden,[22] and researchers from Washington University, USA, isolated the Washington University polyomavirus (WU).[23] KI was isolated from a virus-enriched DNA-cDNA library constructed using nasopharyngeal aspirates,[22] and WU was isolated from respiratory secretions using high-throughput DNA sequencing.[23] KI and WU were found to be closely related, but were divergent from both BK and JC.[23] The genomic organization of KI and WU revealed regions encoding STAg and LTAg (including putative Rb and p53 binding domains).[22, 23] Surprisingly, two partially overlapping LTAg

binding sites in WU were present and had varied spacing when compared to BK and JC.[23] Both KI and WU were found to lack the putative *agno* gene encoded by BK and JC.[22, 23]

Subsequent to the discovery of KI and WU, MCV was identified using digital transcriptome subtraction of cDNA libraries prepared from MCC tissue, and additional examination of the MCV genome revealed mutations and deletions that supported carcinogenesis.[12] Upon further study, clonal integration of the MCV genome was identified in nearly all MCC cells, with the number of integrated copies found to range from a single copy to several thousand.[24, 25, 26] The wild type MCV genome encodes the full length LTAg, although truncation of the STAg in some MCV was found.[27] MCV does not encode the agnoprotein.[12]

An improved RCA technique resulted in the simultaneous isolation of human polyomaviruses 6 and 7 (HPyV6 and HPyV7) from the skin,[27] and TSV from the hair follicle of an immunocompromised patient suffering from TS.[9] Both LTAg and STAg were encoded by the HPyV6 and HPyV7 genomes.[27] TSV genome analysis revealed the LTAg with a putative Rb-binding motif, a sequence previously described to be involved in p53-complex formation, and a protein phosphatase 2A subunit binding motif.[9] On the basis of the anatomical sites where PyV DNA had been isolated, it was suspected that skin cells and tissues supported infection by MCV, HPyV6, HPyV7, and TSV, implying that these viruses may be skin tropic.

Human polyomavirus 9 (HPyV9) was isolated from the serum of a kidney transplant recipient under immunosuppressive treatment.[28] The early coding region contained both LTAg (with a clear Rb-binding domain) and STAg, but sequences for a putative agnoprotein were not found.[28] HPyV9 (or a variant) was assumed to have been re-identified on the skin of MCC patients at a higher frequency than on healthy persons, and may have been shed from the skin.[29]

Finally, human polyomavirus 10 (HPyV10) virus was isolated from skin warts derived from a patient with the rare warts, hypogammaglobulinemia, infections, and myelokathexis (WHIM) syndrome.[30] Sequences for both the LTAg and STAg were identified (although no agnoprotein was found).[30] HPyV10 was found to be nearly identical to a PyV identified in the stool of a child from Malawi, and may represent a new subclade of PyVs collectively referred to as Malawi/human polyomavirus 10.[31, 32, 33]

### 1.1.2 Seroprevalence of polyomaviruses in adults

PyV infections are ubiquitous within human populations (reviewed in DeCaprio & Garcea, 2013),[4] but seroepidemiologic studies have primarily been conducted in European populations or among blood donors (Table 1.1). BK was first isolated from the urine of a renal allograft patient,[21] and JC was cultured from the brain of a PML patient post-mortem.[20] Serologic analyses suggested both types to be prevalent in the adult population, with antibodies against BK in 55-90%[5, 34, 35, 36, 37, 38, 39] and against JC in 20-90% of the study populations.[5, 34, 35, 37, 38, 39] KI[22] and WU[23] were each isolated from the nasopharyngeal aspirates of acute respiratory disease patients, and had reported seroprevalences of 55-90% for KI[3, 5, 14, 40, 41, 42, 43, 44] and 69-98% for WU by adulthood.[5, 14, 40, 42, 43, 44, 45] MCV was found integrated into the genome of an MCC tumor – a rare but aggressive human tumor of the skin characterized by high metastatic potential,[12] and reported MCV seroprevalences among those without MCC ranged from 61-89%.[27, 37, 46, 47, 48, 49]

PyV-linked pathologies in humans rekindled interest in *Polyomaviridae*, leading to the search for other human PyVs. HPyV6 and HPyV7 were isolated from healthy human skin using an RCA technique, and had reported seroprevalences of 69-83%[27, 48] and 35-64%,[27, 48] respectively. The TSV genome was identified from the plucked facial spines of a severely

immunocompromised heart transplant recipient who had developed TS.[9] The range of seroprevalences reported for TSV was considerably narrower than other PyVs, varying from 70-80% in adults,[41, 48, 50] although this virus had been less studied. HPyV9 was amplified from the serum of a kidney transplant recipient,[28] and the literature suggested seroprevalences for HPyV9 ranged from 33-47%.[46, 48] RCA was again used to identify HPyV10 in condyloma specimens from a patient with the rare genetic disorder, WHIM syndrome.[30] HPyV10 seroprevalence was recently reported as 42-66% in adulthood.[51, 52]

Differences in the seroprevalences reported in the peer-reviewed literature could represent true differences between geographical regions, but could also reflect differences in study populations, techniques used for PyV antibody detection and antigen preparation, and definition of seropositive cutoff values. While these studies suggest PyVs are highly prevalent, there is limited information on the distributions of these viruses in the United States population, and potential risk factors for infection are not well-established.

Table 1.1. Selected human polyomavirus (PyV) seroprevalences among healthy, immunocompetent, adults reported in the peer-reviewed literature.*

| Human PyV types | Year of discovery, (reference) | Seroprevalence, (%) | No. seropositive/ No. of study subjects‡ | Remarks† | References |
|---|---|---|---|---|---|
| BK virus, BK | 1971 (Gardner *et al.*, 1971) | 55-90 | Healthy individuals | Numerous serological studies | Reviewed in Knowles, 2006; Moens *et al.*, 2013; |
| John Cunningham virus, JC | 1971 (Padgett *et al.*, 1971) | 20-90 | Healthy individuals | Numerous serological studies | Reviewed in Knowles, 2006; Moens *et al.*, 2013; |
| Karolinska Institutet polyomavirus , KI | 2007 (Allander *et al.*, 2007) | 55-90 | Healthy individuals | Numerous serological studies | Reviewed in Moens *et al.*, 2010 |
| Washington University polyomavirus, WU | 2007 (Gaynor *et al.*, 2007) | 69-98 | Healthy individuals | Numerous serological studies | Reviewed in Moens *et al.*, 2013; Ciotti *et al.*, 2013; |
| Merkel cell polyomavirus, MCV | 2008 (Feng *et al.*, 2008) | 61 | 58/95 | American blood donors (47-75 years of age) | Schowalter *et al.*, 2010 |
| | | 68.6 | 456/665 | General Czech Republic population (≥20 years of age) | Šroller *et al.*, 2013 |
| | | 80 | 38/47 | German volunteers (24-79 years of age) | Pastrana *et al.*, 2012 |
| | | 87 | 461/530 | Italian blood donors (≥20 years of age) | Nicol *et al.*, 2013 |
| | | 88 | 42/48 | American older adults (47-75 years of age) | Pastrana *et al.*, 2009 |
| | | 89.2 | 166/186 | Italian blood donors (18-85 years of age) | Nicol *et al.*, 2012 |
| Human polyomavirus 6, HPyV6 | 2010 (Schowalter *et al.*, 2010) | 69 | 66/95 | American blood donors (47-75 years of age) | Schowalter *et al.*, 2010 |
| | | 83.4 | 442/530 | Italian blood donors (≥20 years of age) | Nicol *et al.*, 2013 |
| Human polyomavirus 7, HPyV7 | 2010 (Schowalter *et al.*, 2010) | 35 | 33/95 | American blood donors (47-75 years of age) | Schowalter *et al.*, 2010 |
| | | 63.6 | 337/530 | Italian blood donors (≥20 years of age) | Nicol *et al.*, 2013 |
| *Trichodysplasia spinulosa*-associated polyomavirus, TSV | 2010 (van der Meijden *et al.*, 2010) | 70 | 104/149 | Finnish university students and staff | Chen *et al.*, 2011 |
| | | 76.4 | 405/530 | Italian blood donors (≥20 years of age) | Nicol *et al.*, 2013 |
| | | 80 | 304/380 | Dutch population-based serum bank (≥20 years of age) | van der Meijden *et al.*, 2011 |
| Human polyomavirus 9, HPyV9 | 2011 (Scuda *et al.*, 2011) | 32.8 | 61/186 | Italian blood donors (18-85 years of age) | Nicol *et al.*, 2012 |
| | | 39.4 | 209/530 | Italian blood donors (≥20 years of age) | Nicol *et al.*, 2013 |
| | | 47 | 154/328 | German adults and adolescents (16-72 years of age) | Trusch *et al.*, 2012 |
| Malawi/human polyomavirus 10, HPyV10 | 2012 (Buck *et al.*, 2012) | 41.8 | 221/528 | Italian blood donors (≥20 years of age) | Nicol *et al.*, 2014 |
| | | 66 | 248/377 | American adults (≥21 years of age) | Berrios *et al.*, 2014 |

* This table is not exhaustive and does not necessarily represent every study conducted to investigate the seroprevalence of human polyomaviruses. It is merely presented to provide a sense as to reasonable PyV seroprevalences.
‡ Some fractions and age groups of interest calculated from % seropositive when not provided directly in the reference.
† Method of seroprevalence ascertainment for polyomavirus has not yet been standardized, although most laboratories use some form of ELISA against VP1.

### 1.1.3 Human polyomavirus serology

Hemagglutination inhibition assays and virus infectivity neutralization assays were the original standard methods for the measurement of antibodies against PyVs, and gradually gave way to the use of enzyme immunoassay technology as the preferred method for measurement. Serological assays have evolved, and the use of recombinant VP1 capsid proteins specific for each viral type has enabled the development of high-throughput screens and sensitive competition assays to distinguish between a history of infection with the various PyVs.

At the time of publication (year 2015), serologic analysis could be performed on the first ten discovered human PyVs: BK,[21] JC,[20] KI,[22] WU,[23] MCV,[12] HPyV6, HPyV7,[27] TSV,[9] HPyV9,[28] and HPyV10.[30] Antigen preparation and techniques used for human PyVs[39, 53, 54] closely followed methods previously described for human papillomaviruses (HPV).[55, 56] Briefly, plasma samples were tested for antibodies against PyV antigens (capsid protein VP1 for BK, JC, KI, WU, MCV isolate 344, HPyV6, HPyV7, TSV, HPyV9, HPyV10; LTAg for BK, JC, MCV, HPyV6, HPyV7, TSV, HPyV10; and STAg for MCV). The multiplex antibody detection approach was based on a glutathione S-transferase (GST) capture enzyme-linked immunosorbent assay (ELISA) method in combination with fluorescent bead technology (Luminex Corp., Austin, Texas, USA, Figure 1.1).[55, 57, 58] Plasma samples – masked to individual identity and characteristics as well as case-control status, were shipped to the German Cancer Research Center (DKFZ; Heidelberg, Germany) on dry ice for PyV serologic analysis by the laboratory that developed the technology.

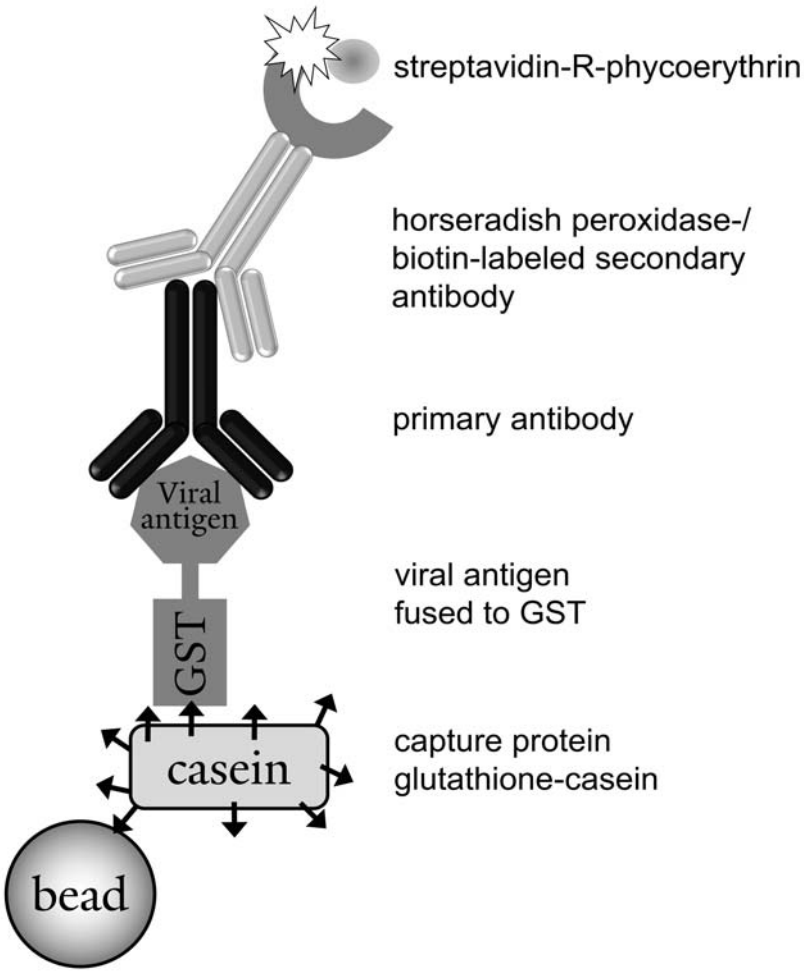



Figure 1.1. Glutathione S-transferase (GST) capture enzyme-linked immunosorbent assay (ELISA) method (adapted from Waterboer *et al.*, 2005). Simplistically, PyV GST fusion proteins are recombinantly expressed in *Escherichia coli*, and indirectly bound to the surface of a bead. Several antigens can be tested simultaneously under identical conditions using an assay platform, and quantities of human immunoglobulin (Ig) G and M (primary antibody) bound to the antigens are detected with a flow cytometer-like analyzer using fluorescent reagents bound to a secondary anti-human IgG/M antibody.

PyV seroreactivity is used as a marker of PyV infection.[39, 59] Seroreactivity against PyV proteins is expressed as the median fluorescence intensity (MFI) of 100+ beads of the same internal color. MFI values reflect antibody affinity, titer, and reactivity, as determined by dilution series.[60] Serum MFI values also correspond to the presence of viral DNA within tumor tissues, with MCV DNA-positive SCC tumors (but not MCV DNA-negative SCC tumors) having significantly higher serum MFI values compared to controls.[54] In addition, MCV viral load and antibody titer were shown to have a strong positive monotonic correlation,[49] although less is known about the other PyV types. The GST capture of recombinantly expressed VP1 capsid proteins was shown to be a reliable technique to assess PyV seroreactivity and has been used as a marker of PyV infection in prior studies,[39, 53, 54, 59] as well as previously being used in serologic studies on HPV infection.[58, 61, 62]

### 1.1.4 Potential role of polyomaviruses in skin carcinogenesis

All human PyVs discovered thus far encode potentially oncogenic proteins that allow them to act as oncoviruses in cell culture and animal models. Indeed, the virus' name is derived from the Greek *poly* meaning 'many' and *oma* referring to 'cancers' observed in infected murine models.[1] The early PyV proteins, LTAg and STAg, provide PyVs with their tumorigenic properties. LTAg contains the LxCxE motif in the Rb-binding pocket which is necessary for transforming activity and cell proliferation,[3, 63] and also harbours a p53-binding domain that binds and inactivates the p53 protein to induce cell division.[3, 16, 63, 64] The J domain of the LTAg cooperates with the LxCxE motif to disrupt the interaction between Rb and the E2F family transcription factors to promote cell cycle entry and progression. Disruption of p53 and Rb functions is necessary but not sufficient to maintain transformation, indicating the requirement of other processes for transformation.[64] STAg contributes to the oncogenic potential of PyVs through the inactivation

of protein phosphatase 2A,[3] which stimulates the MAP kinase pathway and cellular proliferation.[65, 66] The late regions of BK and JC also encode an agnoprotein which may exert its tumorigenic influence through cell cycle dysregulation, interference with DNA repair processes, and chromosome instability.[67, 68, 69] BK, JC, and MCV encode a miRNA that downregulates LTAg expression, which may allow the virus to escape the immune system.[70] Despite mechanistic and experimental evidence for oncogenicity, further research is needed to investigate the relationship between PyVs and human cancers.

Although LTAg is required for viral replication, mediation of oncogenic transformation by PyV requires the separation of TAg expression from the latter parts of the viral life cycle – including DNA replication and host cell lysis.[12] Truncating mutations in TAg are characteristic of MCV DNA sequences amplified from MCC skin tumors following viral genome integration, but the domains required for Rb-induced cell transformation are preserved while viral replication is eliminated (and thus the likelihood of cell survival is increased), and the p53-binding site is removed.[12] It has been suggested that the truncation of LTAg may be a feature of PyV-mediated carcinogenesis,[12] implicating Rb inactivation as the mechanism through which MCV infection drives transformation. These mutations would also suggest MCV is not merely a secondary infection of MCC tumors, as the truncating mutation prevents viral replication.[12] Truncation of the TAg has not been described in other human PyVs.

MCC of the skin is the only human cancer found to contain clonally integrated PyV DNA. MCV has been detected at low levels in other human cancers, including BCC and SCC, although not integrated into the genome.[13, 54, 71] BK persists in the kidneys and urinary tract,[72, 73] which prompted several investigations into possible relationships between BK and urogenital carcinomas (with a tentative association with prostate cancer through loss of LTAg expression

proposed[74, 75]) or bladder cancer (in immunocompromised patients during renal transplantation [76, 77]). Despite the detection of BK and JC in a variety of gastrointestinal and genitourinary cancers, evidence of PyV genome integration is lacking.[78, 79, 80]

The skin tropism of human PyVs and their dermatopathologic potential have not been thoroughly investigated. BK virus DNA has been amplified in 25/33 (76%) of healthy skin tissues, and BK and JC sequences were detected in Kaposi's sarcoma skin lesions.[81] However, several studies decline the skin as a natural site for BK or JC replication. KI and WU have not been isolated from skin samples[27] or melanoma biopsies,[82] suggesting neither virus possesses skin tropism. MCV virions are shed from the skin at high numbers, suggesting that infection is likely not restricted to Merkel cells, and common cells in the epidermis like keratinocytes or melanocytes are involved in the production of virus particles.[27] As with MCV, HPyV6 and HPyV7 virions appear to be shed continuously from the skin (although at a lower number than MCV), and thus are thought to be skin tropic.[27] The TSV genome was first sequenced from the spicules of a young TS patient, but only 3 of the 69 plucked eyebrows collected from immunocompetent renal transplant recipients had evidence of the TSV viral genome,[9] suggesting the occurrence of TSV in skin requires additional research. The occurrence of HPyV9 and HPyV10 in the skin has, at the time of publication, not been thoroughly published on in the literature.

Increased incidence of cancers in immunocompromised patients often suggests a viral etiology due to impaired immunological control over a typically harmless infection, and those who are immunocompromised are at a much higher risk of keratinocyte cancers (KC). [83, 84] Recent epidemiologic studies suggest an oncogenic role for PyV infection in the development of cutaneous squamous cell carcinoma (SCC). In a clinic-based case-control study in Florida, USA,

MCV seropositivity was associated with an increased risk for SCC (OR=1.58, 95% CI: 0.96-2.60), and the association was stronger among those with tumors containing MCV DNA (OR=2.49, 95% CI: 1.03-6.04),[54] although these findings could not be replicated in an expansion of the study group.[71] MCV sequences are also found more frequently in the nonmelanoma skin cancer (NMSC) tumors of immunosuppressed patients when compared to immunocompetent patients, with a 2-fold higher number of MCV positive SCCs in immunosuppressed patients.[85] Conversely, no association was found between SCC and JC seropositivity assessed in samples collected at least 1 month prior to diagnosis in a large Swedish study.[86] Further, a case-control study conducted within organ transplant recipients found no relation between SCC following surgery and the seroprevalence of several types of PyVs (i.e., JC, KI, WU, MCV, HPyV6, and HPyV7) prior to transplantation.[87] However, epidemiologic studies exploring the effect of other PyV infections (apart from MCV) on the risk of SCC are limited, despite evidence of skin tropism for multiple PyVs, including MCV, TSV, HPyV6, and HPyV7.[3]

## 1.2 KERATINOCYTE CARCINOMAS (CUTANEOUS SQUAMOUS CELL CARCINOMA)

Keratinocyte carcinomas include basal (BCC) and squamous cell carcinomas that arise from keratinocytes or their precursors.[88] KCs comprise the most common malignancies in humans,[89] and incidence rates appear to be increasing over time.[90] KCs are often referred to as nonmelanoma skin cancers, which can include rarer forms of skin cancer.[88, 89, 91, 92] SCC is subclassified into the following histologic types: Squamous Cell Acantholytic, Spindle-cell, Verrucous, and Adenosquamous cell; with *in situ* disease classified as SCC *in situ* (Bowen's disease).[93] Intraepidermal proliferations (dysplasias), actinic keratosis (AK), and SCC *in situ*, may represent precursor lesions of SCC.[94] The tumor progression model for SCC is still being

studied, but is thought to involve multiple steps which may include *TP53* alteration induced by ultraviolet radiation (UVR),[95, 96, 97] Src-family kinases (*SFK*),[98] *Ras* oncogene,[99] and activating transcription factor 3 (*ATF3*).[100]

### 1.2.1 Disease burden

Cutaneous squamous cell carcinoma arises from epithelial keratinocytes, with increasing incidence rates reported in the United States,[101, 102, 103, 104] Australia,[105, 106] and Europe.[107, 108] In the United States, an estimated 186,157-419,543 individuals were diagnosed with SCC of the skin, and 3,932-8,791 SCC patients died from their disease, in the year 2012.[102] Although the prognosis is generally favorable, the disfigurement and morbidity resulting from SCC ranks as an important public health concern and under-recognized health burden.

The geographic and sociodemographic features of SCC cases have remained consistent over the past several decades, with a gradient of increasing incidence rates with proximity to the equator, impacting primarily fair-skinned populations, and increasing with age.[104] Men generally have 2-4 times the incidence of SCC when compared to women.[91] Individuals with higher socioeconomic status (SES) living in the least deprived regions (e.g., professional workers, greater educational attainment, car ownership) may have increased risk of KCs, but this research has primarily been conducted in northern European countries,[109, 110] and may be attributed to leisure time spent in sunnier climes.[110]

Efforts to document the incidence of SCC around the world have greatly expanded (reviewed in Lomas *et al.*, 2012),[90] although the standardization of rates to different standard populations makes direct comparisons of incidence difficult. SCC studies have primarily been conducted in white populations from the USA, Europe, and Australia. Worldwide incidence rates for SCCs vary over 100-fold in white populations (Figure 1.2). Australia continues to report the

highest rates in the world, with an average incidence rate for SCC of 387 per 100,000 person-years in 2002 age-standardized to the world standard population.[111] On continental Europe, Italy and Switzerland appear to have some of the highest SCC incidence rates.[112, 113, 114, 115, 116, 117] The high incidence rates of SCC reported in Switzerland may be attributed to having the highest average altitude in Europe.[90] The average incidence rate in England for SCC has been calculated to be 22.65 per 100,000 person-years in the 2000s.[90] The lowest reported incidence rates among populations of European descent come from Croatia, with rates for SCC of <10 per 100,000 person-years between 2003 and 2005, although these estimates were standardized to the 2001 Croatian census population.[118]

In the United States, Southern state Arizona has reported some of the highest SCC incidence rates of >100 per 100,000 person-years standardized to the 1970 USA population.[119] In contrast, Northern states New Hampshire and Minnesota have SCC rates of approximately 30 per 100,000 person-years[120, 121, 122] – comparable to those found in Canada.[123, 124, 125]

The lowest rates of SCC occur in populations of largely non-European descent (e.g., India, Asia, and Africa).[126, 127, 128] The limited data on non-white populations in the United States also suggest racial or ethnic variation in KC incidence rates. For example, In Arizona, USA, SCC rates among Hispanics were 13.8 and 32.9 per 100,000 person-years among women and men respectively, while the rates were 11-fold higher in non-Hispanic whites.[119] Despite lying at a similar latitude to North America, a study conducted in Jordan (consisting of darker-skinned people of Arabian descent) using the Jordan Cancer Registry from 1997 to 2001 found SCC incidence rates of 14.2 for men and 4.18 for women per 100,000 person-years standardized to the USA population for the years 1997-2000.[129] Among Asians in Singapore, SCC rates were over 2-fold higher among the lighter-skinned Chinese than the Malays and Indians.[130]

Interestingly, a recent study conducted in a Southern region of Brazil (with a higher proportion of those of European decent) found the incidence of invasive or *in situ* SCC to be 94.39 per 100,000 person-years in 2006.[131] While lack of standardized reporting methods and ethnic differences complicate international comparisons, available data from the USA and Australia indicate a distinct gradient of increasing incidence rates with proximity to the equator. [111, 132, 133, 134, 135, 136]

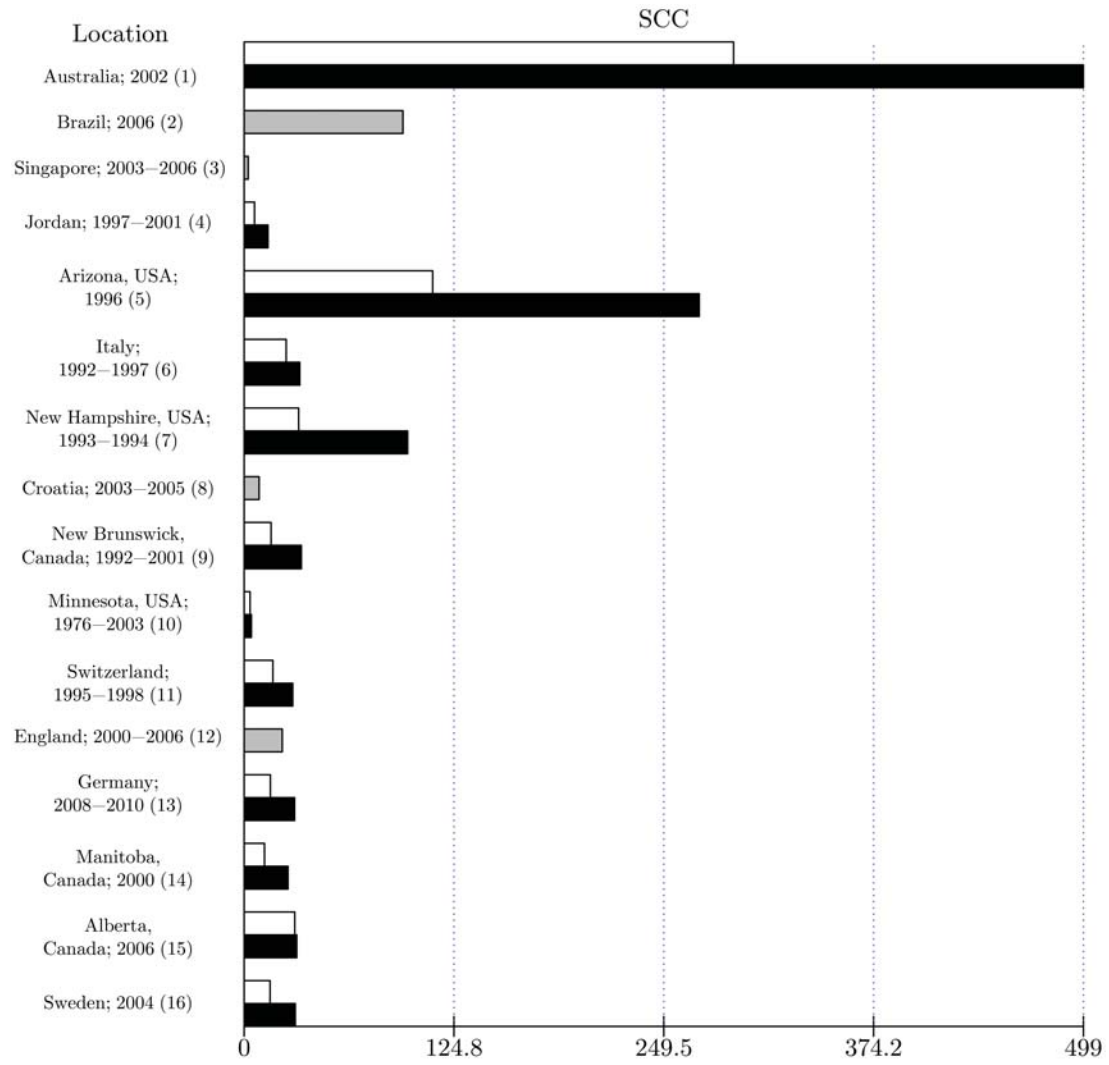



Figure 1.2. Selected incidence rates (per 100,000 person-years) for cutaneous squamous cell carcinoma (SCC) worldwide. Male rates are black, female rates are white, and rates including both males and females are gray. The rate values presented for the corresponding time periods were drawn from the following references: (1) Staples *et al.*, 2006, (2) Nunes *et al.*, 2009 a population of primarily European ancestry, (3) Sng *et al.*, 2009 a Malay, Asian, and Indian population, (4) Omari *et al.*, 2006 a population of primarily Arabian descent, (5) Harris *et al.*, 2001, (6) Boi *et al.*, 2003, (7) Karagas *et al.*, 1999, (8) Lipozenčić *et al.*, 2009, (9) Hayes *et al.*, 2007, (10) Christenson *et al.*, 2005, (11) Levi *et al.*, 2001, (12) Lomas *et al.*, 2012, (13) Eisemann *et al.*, 2014, (14) Demers *et al.*, 2005, (15) Jung *et al.*, 2010, (16) Dal *et al.*, 2008. The populations from references (1, 2, and 5-16) are primarily of European ancestry; the populations from references (3-4) are of non-European ancestry. The references (1, 6, 9, 10, 11, and 14) were world-standardized, (12, 13, and 16) were European-standardized, (3, 4, 5, 7, 8, and 15) were standardized to the population structure of that country, and (2) was not standardized to a population outside of the study sample.

A number of studies indicate a dramatic worldwide increase in the number of cases of KC (reviewed in Lomas *et al.*, 2012),[90] and the incidence rates of KCs may be increasing more rapidly than any other malignancy.[137] For SCC, USA surveys done in the 1970s found a 1-2% average increase per year in age-adjusted incidence rates.[132] Increases in SCC rates now range from 3-19% each year in the USA and Canada,[120, 123, 136, 138, 139, 140] with the exception of a decline in invasive SCC in Arizona.[119] In Europe, the estimated average annual percentage change (AAPC) for SCC rates was 1-5% in Scotland,[141, 142] and 10-13% in the UK,[140] for different time periods since the late 1970s. Studies of the Nordic countries have found SCC incidences in Sweden, Norway, Denmark, and Finland to have also increased annually by 3-12% since the 1960s and 1970s to more recent years.[143, 144, 145, 146] The disease burden for SCC in the Netherlands increased by 2-10% annually since the 1980s with no evidence of plateauing in the number of cases.[147, 148, 149, 150, 151] Other regions of the world also report AAPC increases over time such as Australia (reviewed in Perera *et al.*, 2015)[152] with evidence that the rate of increase has been slowing,[90, 111] in Serbia,[153] and in Iran.[154] The reasons for increasing incidence rates are not fully understood, but may be related to stratospheric depletion of the world's ozone layers increasing the amount of ambient UVR,[155, 156] or an increase in sun exposure behavior (e.g., leisure time outdoor activities). Enhanced or earlier detection of SCC cases are unlikely to account for these temporal trends.[116, 157, 158, 159, 160, 161]

There is evidence that the worldwide mortality rates of KC are declining slightly,[89] but survival rates for SCC have been relatively stable.[149] Five-year relative survival after SCC has been estimated to be 90%,[162] which is consistent with clinical observations that a relatively small subset of patients develop metastatic disease that can be fatal.[163] For patients who have SCC metastases, the 10-year survival is less than 20% for patients with regional lymph node

involvement, and less than 10% with distant metastases.[164] Men have a higher rate of mortality from SCC than women.[165] Independent predictors of poor outcomes include a tumor diameter at least 2 centimeters, invasion beyond fat, poor differentiation, ear or temple location, perineural invasion, and anogenital invasion.[163, 166] SCC patients experience higher overall mortality not attributable to smoking history nor could be explained by their heightened risk of other malignancies or intervening cancers,[167] but may be due to other factors such as alcohol use or immunosuppression.[165, 168, 169] Although fatal in only a small fraction of SCC cases, certain subgroups of the population experience appreciable mortality from these tumors, such as the immunocompromised (e.g., patients who have undergone an organ transplantation or who have HIV/AIDS).[104]

### 1.2.2 Etiologic factors

The skin is the primary barrier against the external environment, and epidermal cells are often the first cells to be exposed to potentially carcinogenic agents. While the etiology of cutaneous SCC is likely multifactorial including both environmental and endogenous risk factors, the exact mechanisms by which risk factors cause SCC are not fully understood (Figure 1.3). Established risk factors include ultraviolet light exposure,[96, 170] skin pigmentation,[171, 172] exposure to ionizing radiation,[173, 174] arsenic,[175, 176, 177, 178] and polycyclic aromatic hydrocarbons (PAH).[179, 180] Chronic ultraviolet radiation[91] and artificial UVR exposure (e.g., indoor tanning lamps,[181, 182, 183, 184, 185] psoralen UVA (PUVA) therapy for psoriasis[186, 187, 188]) are consistently associated with an elevated risk of SCC through the induction of DNA damage[189, 190, 191] and immunosuppression.[192] Photodamage is recognizable clinically, as chronic sun exposure results in skin photoaging that progressively destroys the organized basket weave pattern and overall elasticity of the skin.[193, 194, 195, 196, 197] Light pigmentary traits for hair and eye color (e.g., fair

skin, blue eyes, and blonde or red hair) are accepted risk factors for KCs.[198, 199, 200] Investigations have also documented an increased risk of SCC among those with a freckling phenotype[201, 202] and nevi.[135, 198, 203] Elevated risks of SCC following exposure to ionizing radiation have been observed in several occupationally exposed groups, including uranium miners and quarrymen,[204] early medical radiation workers,[205] and radiologic technologists.[206] Occupational exposure to metalworking fluids that often contain PAHs, such as among autoworkers, has been consistently linked to SCC (particularly of the scrotum).[207] A prior SCC diagnosis is a strong predictor of a subsequent SCC.[208, 209, 210] In addition, patients with autosomal recessive genetic defects in the ability to correct DNA photolesions (e.g., xeroderma pigmentosum (XP)),[211, 212] and those who are immunocompromised (e.g., transplant recipients or those given immunosuppressive medications),[84, 213, 214] carry a high risk of SCC.

Emerging risk factors requiring additional study confirmation include the use of BRAF inhibitors,[215] photosensitizing drugs (e.g., tetracyclines, fluoroquinolones, thiazide diuretics),[216, 217, 218, 219, 220] sex steroids (particularly high doses of oral contraceptives[221, 222, 223, 224, 225, 226] and hormone replacement therapy containing high levels of exogenous estrogen),[225, 227] and certain analgesics (e.g., aspirin) and nonsteroidal anti-inflammatory drugs (NSAIDs)[228, 229, 230, 231, 232] that act through a *COX-2* pathway for a chemopreventive effect.[232] Tumorigenesis may be enhanced by a hyperactive immune system and chronic inflammation, or alternatively could suppress cancer cells due to heightened immune surveillance, which has lead to increased exploration of the association between atopic and allergic states[233, 234] in the etiology of KCs resulting in some suggestions of an increased SCC risk among asthmatics[235] and those with a history of atopic dermatitis.[236, 237] An increased risk of SCC occurrence with higher dietary fat intake has been found,[238, 239, 240] and observational data regarding the consumption of various

fatty acids (e.g., omega-3 and omega-6 polyunsaturated fatty acids (PUFAs)[241, 242, 243, 244, 245, 246]), leafy vegetables,[239, 240, 247] alcohol use,[248, 249] tea consumption,[244, 250, 251, 252, 253] caffeine,[237, 254, 255, 256] and other micronutrients such as vitamins D,[257, 258] B3,[259, 260] and A[261, 262, 263, 264, 265] are growing. Other dietary factors with potential photoprotective effects (e.g., garlic, curcumin, lycopene, limonene) are being investigated, but have not been assessed in trials. Tobacco smoke, a source of PAH exposure[266] and reduced antioxidant defenses,[267, 268, 269] has been associated with an increased risk of SCC – particularly with current smoking.[91, 270, 271, 272, 273, 274]

Organ transplant recipients who require potent immunosuppressive therapy to prevent allograft rejection experience extreme risks of NMSC.[83, 275] In addition to increased incidence of skin cancer in transplant recipients, SCCs appear to be more aggressive with higher rates of recurrence and metastasis.[275] Increased incidence of cancers in immunocompromised patients often suggests a viral etiology. SCC tumors from epidermodysplasia verruciformis (EV) patients were found to contain HPV DNA, and this viral DNA has been detected in up to 90% of SCC tumors in EV patients (who are already extremely susceptible to skin cancer).[276, 277] Thus far, genus β HPVs[61, 278] have been associated with an increased risk of SCC, raising the possibility of other viral risk factors for the development of this skin cancer, and an etiologic role for PyVs has been recently raised in the peer-reviewed literature.[54, 71, 279]

Figure 1.3. Exogenous factors related to cutaneous squamous cell carcinoma (SCC).

| Exposure | SCC |
|---|---|
| *Environmental Exposures* | |
| Total lifetime sun exposure (chronic) | ●● |
| Childhood sun exposure | ● |
| Recent sun exposure | ● |
| Indoor tanning (non-medical) | ●● |
| PUVA | ●● |
| Ionizing radiation | ■ |
| Arsenic | ●● |
| PAH (occupational) | ●● |
| Cigarette smoking | ● |
| *Medications* | |
| Immunosuppressive therapy | ●● |
| BRAF inhibitors | ▲ |
| Photosensitizing medications | ■■ |
| NSAIDS | □ |
| Sex steroids | ■ |
| *Viruses* | |
| Human papillomavirus infection | ■ |
| Human polyomavirus infection | ▲ |
| *Dietary Factors* | |
| Caffeine | △ |
| Tea | △ |
| Vitamin D | △ |
| Vitamin B3 | △ |
| Vitamin A | △ |
| High fat diet | ▲ |
| Omega-3 to Omega-6 ratio | △ |
| Omega-6 | ▲△ |

Legend: ● filled circle = well established risk factor; ○ empty circle = well established protective risk factor; ■ filled square = possible risk factor; □ empty square = possible protective risk factor; ▲ filled triangle = exploratory risk factor; △ empty triangle = exploratory protective risk factor. Two circles/squares/triangles reflects a stronger association in the specified direction.

## 1.3 THESIS AIMS

The following thesis addresses critical gaps in our understanding of the risk factors for PyV infection and its relation to skin cancer occurrence within the United States. We first aimed to determine the seroprevalence and risk factors for seropositivity of human polyomavirus types among adults in a general United States population. Next, we investigated the association between human polyomavirus infection and risk of cutaneous squamous cell carcinoma in the New Hampshire Skin Cancer Study. Lastly, we performed a prospective study of human polyomaviruses and risk of cutaneous squamous cell carcinoma in the Skin Cancer Prevention Study conducted in the United States.

# Chapter 2

# Seroepidemiology of human polyomaviruses in a United States population

*Published as: Anala Gossai, Tim Waterboer, Heather H. Nelson, Angelika Michel, Martina Willhauck-Fleckenstein, Shohreh F. Farzan, Anne G. Hoen, Brock C. Christensen, Karl T. Kelsey, Carmen J. Marsit, Michael Pawlita, and Margaret R. Karagas. "Seroepidemiology of Human Polyomaviruses in a US Population." American Journal of Epidemiology 183 (2016): 61-69. (excluding sections 2.3.5, 2.4.4, & **bold italicized** paragraphs of 2.5).*

## 2.1 ABSTRACT

Polyomaviruses (PyV) are potentially tumorigenic viruses in humans. However, limited data exists on the population seroprevalence of PyVs and individual characteristics that relate to seropositivity. Using multiplex serology, we determined the seroprevalence of ten human PyVs (BK, JC, KI, WU, MCV, HPyV6, HPyV7, TSV, HPyV9, and HPyV10) among controls from a population-based skin cancer case-control study (n=460). On a subset of participants (n=194), methylation at CpG sites across the genome was measured in peripheral blood using the Illumina Infinium HumanMethylation27 BeadChip array, from which lymphocyte subtype proportions were inferred. All participants were seropositive for at least one PyV, with seroprevalences ranging from 17.6% (HPyV9) to 99.1% (HPyV10). Seropositivity to JC, MCV, and HPyV7 increased with age. JC and TSV seropositivity was more common among men than women. Smokers were more likely to be HPyV9 seropositive but MCV seronegative, and HPyV7 seropositivity was associated with prolonged glucocorticoid use. In peripheral blood, differential methylation at CpG sites depending upon PyV seroreactivity did not reach statistical significance at the FDR level ($q$<0.05), but genes associated with cancer pathways and the immune response were over-represented in a gene ontology analysis of CpG sites with $P$<0.05. Cancer and immune pathways were over-represented in an exploratory gene ontology analysis of CpG sites differentially methylated when PyV seropositive. Based on DNA methylation profiles, differences were observed in CD8+ T and B cell proportions by BK, JC, and HPyV9 seropositivity. Our findings suggest seropositivity to PyVs is common in the United States, and varies by sociodemographic and biologic characteristics, including those related to immune function.

## 2.2 INTRODUCTION

Polyomaviruses (PyV) are non-enveloped viruses with an icosahedral capsid ~45 nm in diameter, containing a circular double-stranded DNA genome.[2, 3] The PyV genome contains genes encoding three capsid proteins (VP1, VP2, VP3), as well as small and large T antigens (TAg).[2] Members of the *Polyomaviridae* family have been discovered in humans, nonhuman primates, mice, birds, bats, and a host of other species.[1] Urine or other bodily fluids have been implicated as the main vehicle for transmission within families or intimate contacts.[280] However, the diversity of the human polyomavirome and its impact on human health has not yet been fully elucidated, reflected in the continued discovery of new PyVs.

Although largely based on studies in European populations and blood donors, PyVs appear to be ubiquitous within human populations (reviewed in DeCaprio & Garcea, 2013).[4] While often asymptomatic, under conditions of immunosuppression, BK virus[21] has been associated with polyomavirus-associated nephropathy (PVAN) and cystitis,[7, 8] JC virus[20] with progressive multifocal leukoencephalopathy (PML),[7, 8] and *Trichodysplasia spinulosa*-associated polyomavirus (TSV) with the skin disease *Trichodysplasia spinulosa* (TS).[9] Furthermore, Merkel cell polyomavirus (MCV) with mutations in the large TAg[12] has been implicated as a causal factor for Merkel cell carcinomas (MCC)[24] regardless of immune status.[13, 14] Although prevalent in various tissues across multiple populations,[48] Karolinska Institute polyomavirus (KI),[22] Washington University polyomavirus (WU),[23] human polyomaviruses 6 and 7 (HPyV6 and HPyV7),[27] human polyomavirus 9 (HPyV9),[28] and Malawi/human polyomavirus 10 (HPyV10)[30] have not yet been associated with any specific disease phenotype.[6]

To determine the seroprevalence of these viruses, and the individual sociodemographic and biologic characteristics associated with seropositivity, we measured the frequency of serum

antibodies against ten PyVs amongst controls from a United States population-based case-control study.

## 2.3 METHODS

### 2.3.1 Study population

The study participants and methods have been described in detail elsewhere.[62, 120] Briefly, our study included controls from a population-based case-control study of basal cell and squamous cell skin cancers. Controls were frequency-matched to the age (25-74 years) and gender distribution of skin cancer cases diagnosed from July 1993 through June 1995. Residents of New Hampshire, USA, were selected from the New Hampshire Department of Transportation (<65 years) and the Center for Medicaid and Medicare Services enrollment lists (≥65 years), and were required to speak English and have a listed telephone number. Sociodemographic information (e.g., age, gender, and level of education), lifestyle factors (e.g., cigarette smoking), sunlight-related characteristics (e.g., response to first exposure in summer to one hour of sunlight, number of severe sunburns, and skin color), and medical history (e.g., use of glucocorticoids for one month or longer, or organ transplant recipient) were collected through personal interviews. All participants provided informed consent in accordance with the Committee for the Protection of Human Subjects at Dartmouth College.

### 2.3.2 Blood sample collection

Venous blood samples of 20-30 mL were collected in heparinized tubes for serologic analysis (as described in Karagas *et al.*, 2006).[62] Blood was separated by centrifugation at 2,500*g* for 20 minutes at 4°C. Each component (plasma, red blood cells, buffy coat) was labeled and stored separately at -80°C until analysis. DNA was also extracted from the buffy coat for DNA methylation analysis. Plasma samples, masked to individual identity and characteristics, were

shipped to the German Cancer Research Center (DKFZ; Heidelberg, Germany) on dry ice for serologic analysis.

### 2.3.3 Human polyomavirus serology

Antigen preparation and techniques used for human PyVs[39, 53, 54] closely follow methods previously described for human papillomaviruses (HPVs).[55, 56] Briefly, plasma samples were tested for antibodies against ten human PyVs (capsid protein VP1 for BK, JC, KI, WU, MCV isolate 344, HPyV6, HPyV7, TSV, HPyV9, HPyV10; large TAg for BK, JC, MCV, HPyV6, HPyV7, TSV, HPyV10; and small TAg for MCV). Monkey viruses (e.g., chimpanzee polyomavirus (ChPyV), simian virus 40 (SV40), and lymphotropic polyomavirus (LPV)) were not considered. The multiplex antibody detection approach was based on a glutathione S-transferase (GST) capture enzyme-linked immunosorbent assay (ELISA) method in combination with fluorescent bead technology (Luminex Corp., Austin, Texas, USA).[55, 57, 58]

Seroreactivity against PyV proteins was expressed as the median fluorescence intensity (MFI) of 100+ beads of the same internal color. MFI values reflect antibody affinity, titer, and reactivity determined by dilution series.[60] Standard cut points for seropositivity were chosen for the MFI of each PyV tested by visual inspection of frequency distribution curves (percentile plots), as done previously;[14, 39, 56, 62] stringent criteria were chosen to increase specificity. For VP1, the cutoff value was 250 MFI units for all ten PyVs (as used in Teras *et al.*, 2014).[281] The cutoff value for TAg was 400 MFI units for BK, JC, HPyV6, HPyV7, TSV, HPyV10, and MCV large TAg; the cutoff value for MCV small TAg was 50 MFI units.

To evaluate the robustness of PyV VP1 seroprevalence estimates, we used a sliding cut point between 50 and 450 MFI units (Supplemental Figure 2.1). We also calculated cut points using a method adapted from van der Meijden *et al.*, 2013.[282] This involved a frequency

distribution analysis with a bin width of 250 MFI units, and the seronegative population defined by either all the bins containing >10% of the study participants, or the first bin. Cutoffs were then calculated as the mean seroresponse for those who were seronegative plus three times the standard deviation. This resulted in similar seroprevalence estimates as the standard cut points, except for JC due to the use of two bins (as opposed to one) in cutoff determination. Given the insensitivity of PyV seroprevalence to cut point definitions, we ultimately used the standard cut points in all analyses.

### 2.3.4 Statistical analyses

We first constructed a phylogenetic tree relating human PyVs from the complete genome sequences stored in the RefSeq database using Muscle 3.8.31 to create a neighbour-joining tree without distance corrections.[283] The accession numbers of the complete genome sequences for the ten assayed human PyVs selected from RefSeq were: NC_001538.1 (BK), NC_001699.1 (JC), NC_009238.1 (KI), NC_009539.1 (WU), NC_010277.1 (MCV), NC_014406.1 (HPyV6), NC_014407.1 (HPyV7), NC_014361.1 (TSV), NC_015150.1 (HPyV9), and NC_018102.1 (HPyV10).

The frequency of PyV seropositivity for each virus was then examined for both PyV seropositivity overall (any PyV positive versus no PyV positive), and by the number of PyV types to which an individual tested positive, using binary MFI cut points. In addition, we used the continuous MFI values to compute Spearman rank correlation coefficients ($\rho$) between each of the PyVs assayed.

We tested the association between various individual characteristics in relation to PyV seropositivity using the $X^2$ test (for categorical variables, i.e., age groups 25-44, 45-54, 55-64, 65-69, and 70-75 years; gender, education, smoking status, skin sun sensitivity, number of

sunburns 0-1 or 2+, and glucocorticoid use), or Fisher's exact test (for categorical variables with small strata, i.e., <10 participants), and Wilcoxon rank sum test or Kruskal-Wallis test (for continuous variables, i.e., age and for the mean number of PyVs seropositive). We further used logistic regression to assess the association between PyV seropositivity and each characteristic of interest while considering the potential confounding effects of other characteristics. Variables that altered associations between the characteristics of interest and PyV seropositivity (i.e., changed our effect estimates by >10% [284]) were included in our model. Only age group and gender were found to change the estimates, so *P*-values adjusted for these variables are presented. All statistical tests were two-sided, and significance was assessed at the $\alpha$=0.05 level. All analyses of the serology data were performed in R version 3.0.1.

### 2.3.5 Analysis of DNA methylation and immune cell proportions

DNA methylation data and inferred immune cell proportions have been previously reported in Marsit *et al.*, 2011 [285] and Koestler *et al.*, 2012,[286] respectively. Briefly, DNA was extracted from peripheral blood using QIAmp DNA mini kit (Qiagen, California, USA), and underwent sodium bisulfite modification using EZ DNA Methylation kit (Zymo, California, USA) according to the manufacturers' protocols. DNA methylation data was obtained using the Illumina Infinium HumanMethylation27 array, using methods described in detail elsewhere.[285, 287, 288, 289] Array quality assurance was assessed and poor-performing loci were removed.[287, 288, 290] Sex-linked loci, SNP-containing CpG loci and CpGs assayed with a SNP-containing probe, and autosomal probes reported to be cross-reactive or repetitive, were removed.[287, 291, 292] The resulting methylation dataset consisted of 20,046 CpG loci (associated with 12,348 genes) for analysis. From methylated (M) and unmethylated (U) allele intensity, $\beta$-values representing the methylation status at each CpG locus were calculated, where the ratio of fluorescent signals

β=Max(M,0)/[Max(M,0)+Max(U,0)+100].[287] The average probe intensity for the M and U alleles was used.[287] DNA methylation values were obtained in the form of β-values ranging from [0,1] with a β distribution.[289] Complete methylation was indicated by β-values near 1 and absence of methylation indicated by β-values near 0.[288, 289] β-values were logit-transformed to M-values [$\log_2(\beta/1-\beta)$] prior to analysis.[287, 289]

The proportion of immune cell subtypes was inferred from DNA methylation profiles using an algorithm described in Houseman *et al.*, 2012.[293] We compared the median proportion of adaptive lymphocyte predictions for CD8+ T cells, CD4+ T cells, and B cells between those who were PyV seropositive versus seronegative for each PyV type except those with <10 participants seronegative (i.e., WU and HPyV10) using the Wilcoxon rank sum test.

The extent of methylation at individual CpG sites (predictor) and PyV seropositivity (outcome) in a locus-by-locus analysis for each of the ten human PyVs, as well as seropositivity for </≥7 (median) PyVs, was examined using logistic regression models. False discovery rate (FDR) was used to adjust for multiple comparisons using the qvalue package in R and a significance threshold of $q<0.05$.[294, 295, 296] All models included age (continuous) and gender as potential confounders.[297, 298] To adjust for potential differences in cell proportions among participants, we included immune cell proportions for CD8+ T cells, CD4+ T cells, B cells, natural killer (NK) cells, monocytes, and granulocytes as continuous variables in our logistic regression models.[299, 300] All tests were two sided, and analyses for the methylation and immune cell proportions data were carried out in R version 3.1.0.

Finally, the official gene symbols of the PyV-associated differentially methylated CpG sites (where $P<0.05$ following adjustment for age, gender, and immune cell proportions) were assessed for gene enrichment and functional annotation with the Database for Annotation,

Visualization and Integrated Discovery (DAVID) version 6.7 (http://david.abcc.ncifcrf.gov/).[301, 302, 303] DAVID consists of an integrated database and analytic tools intending to mine biological meaning from a large list of genes.[301, 302] Gene lists for each PyV were compared with a background of 11,123 genes (90.1% of 12,348 genes associated with the final methylation dataset);[304] and >80% of the differentially methylated genes of interest were found in the DAVID database for *Homo sapiens*. Gene Functional Classifications with an FDR $q<0.05$ were considered significantly enriched after using the highest classification stringency and default parameters.

## 2.4 RESULTS

### 2.4.1 Study population characteristics

Venous blood samples were available for serologic analyses from 460 study participants (85.2% of the 540 interviewed; 1 participant was excluded due to insufficient bead count during serologic analysis). DNA was extracted from the buffy coat samples, and a subset of 194 study participants (42.2% of the 460 with blood) were randomly selected for DNA methylation analysis. No appreciable differences were noted in the characteristics of individuals for whom we did not obtain serology data (data not shown). There were also no notable differences in the characteristics of the 460 individuals with PyV serology data (Table 2.1) and the subset of 194 study participants with paired methylation data (Supplemental Table 2.1).

The study population included a higher proportion of men than women, >50% had skin that tanned following sun exposure, almost half had completed at least high or technical school, and most were former smokers. The mean age of participants was 61.2 years (standard deviation=10.7 years). There was one organ transplant recipient (0.2%), and other participants (8.5%) reported the use of oral glucocorticoids for 1 month or longer (Table 2.1).

Table 2.1. Distribution of human polyomavirus (PyV) seropositivity among 460 study participants by sociodemographic factors. [a]

| Variable | Total, No. (%) [d] | No. PyV types seropositive, mean (SD) [e] | Seropositive, No. (%) [b,c] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | BK | JC | KI | WU | MCV | HPyV6 | HPyV7 | TSV | HPyV9 | HPyV10 |
| Gender | | | | | | | | | | | | |
| Male | 280 (100) | 7.5 (1.5)*** | 246 (87.8) | 168 (60.0)* | 256 (91.4) | 271 (96.8) | 204 (72.8) | 212 (75.7) | 164 (58.6) | 239 (85.4)*** | 54 (19.3) | 277 (98.9) |
| Female | 180 (100) | 7.0 (1.3) | 157 (87.2) | 88 (48.9) | 164 (91.1) | 177 (98.3) | 120 (66.7) | 127 (70.6) | 95 (52.8) | 133 (73.9) | 27 (15.0) | 179 (99.4) |
| Education | | | | | | | | | | | | |
| Elementary to high or technical school | 227 (100) | 7.4 (1.4) | 201 (88.5) | 126 (55.5) | 211 (93.0) | 221 (97.4) | 163 (71.8) | 169 (74.4) | 131 (57.7) | 193 (85.0) | 41 (18.1) | 226 (99.6) |
| Any college | 144 (100) | 7.1 (1.4) | 123 (85.4) | 78 (54.2) | 128 (88.9) | 140 (97.2) | 99 (68.8) | 102 (70.8) | 77 (53.5) | 110 (76.4) | 18 (12.5) | 142 (98.6) |
| Graduate or professional school | 89 (100) | 7.4 (1.6) | 79 (88.8) | 52 (58.4) | 81 (91.0) | 87 (97.8) | 62 (69.7) | 68 (76.4) | 51 (57.3) | 69 (77.5) | 22 (24.7) | 88 (98.9) |
| Smoking status | | | | | | | | | | | | |
| Never | 146 (100) | 7.3 (1.3) | 130 (89.0) | 81 (55.5) | 133 (91.1) | 145 (99.3) | 111 (76.0)*[,h] | 103 (70.5) | 79 (54.1) | 119 (81.5) | 15 (10.3)*[,h] | 144 (98.6) |
| Former | 230 (100) | 7.4 (1.5) | 205 (89.1) | 128 (55.6) | 210 (91.3) | 222 (96.5) | 159 (69.1) | 178 (77.4) | 138 (60.0) | 191 (83.0) | 49 (21.3) | 228 (99.1) |
| Current | 84 (100) | 7.0 (1.6) | 68 (81.1) | 47 (56.0) | 77 (91.7) | 81 (96.4) | 54 (64.3) | 58 (69.0) | 42 (50.0) | 62 (73.8) | 17 (20.2) | 84 (100) |
| Skin sun sensitivity [f] | | | | | | | | | | | | |
| Burn | 144 (100) | 7.1 (1.5) | 124 (86.1) | 77 (53.5) | 128 (88.9) | 141 (97.9) | 97 (67.4) | 105 (72.9) | 71 (49.3) | 119 (82.6) | 20 (13.9) | 142 (98.6) |
| Tan | 314 (100) | 7.4 (1.4) | 277 (88.2) | 179 (57.0) | 290 (92.4) | 305 (97.1) | 225 (71.6) | 233 (74.2) | 187 (59.6) | 251 (79.9) | 60 (19.1) | 312 (99.4) |
| No. of lifetime painful sunburns | | | | | | | | | | | | |
| 0-1 | 235 (100) | 7.4 (1.3) | 206 (87.6) | 136 (57.9) | 212 (90.2) | 231 (98.3) | 172 (73.2) | 176 (74.9) | 135 (57.4) | 187 (79.6) | 40 (17.0) | 234 (99.6) |
| 2+ | 220 (100) | 7.2 (1.6) | 192 (87.3) | 116 (52.7) | 204 (92.7) | 212 (96.4) | 149 (67.7) | 161 (73.2) | 122 (55.4) | 180 (81.8) | 38 (17.3) | 217 (98.6) |
| Glucocorticoid use [g] | | | | | | | | | | | | |
| Yes | 39 (100) | 7.4 (1.3) | 35 (89.7) | 22 (56.4) | 36 (92.3) | 37 (94.9) | 23 (59.0) | 30 (76.9) | 30 (76.9)***[,h] | 32 (82.0) | 5 (12.8) | 39 (100) |
| No | 415 (100) | 7.3 (1.4) | 363 (87.5) | 231 (55.7) | 278 (67.0) | 406 (97.8) | 296 (71.3) | 305 (73.5) | 226 (54.4) | 334 (80.5) | 74 (17.8) | 411 (99.0) |

[a] Numbers may not sum to the overall total due to missing data.
[b] Percentages indicate the proportion of healthy adults who are PyV seropositive versus PyV seronegative in each given strata.
[c] PyV infection was determined using seropositivity for the VP1 protein.
[d] All participants were seropositive for at least one of the ten PyVs assayed.
[e] Standard deviation referred to as "SD".
[f] Sun sensitivity was defined as the reaction to 1 hour of sun exposure the first time in the summer. "Burn" refers to those who experienced severe sunburn with blistering or a painful sunburn and then peeling. "Tan" refers to those who experienced mild sunburn and then some tanning or tanned without sunburn.
[g] Glucocorticoid use was defined as having used oral steroid or corticosteroid medications (for example, cortisone or prednisone) for one month or longer.
[h] Age group and gender-adjusted *P*-value, as determined by logistic regression.
* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ to test difference in proportions between groups, as determined by $X^2$ or Fisher's exact tests for categorical variables, and by Kruskal-Wallis or Wilcoxon rank sum tests for continuous variables.

### 2.4.2 Seropositivity to polyomavirus

We did not find correlations or evidence of cross-reactivity between the VP1 capsid proteins of most PyV types. The exceptions were positive correlations between the closely related HPyV6 and HPyV7 (ρ=0.59, *P*=1.1 e-67), and MCV and HPyV9 (ρ=0.43, *P*=1.0 e-33) (Figure 2.1). However, when restricted to participants who were seropositive for each PyV pair, we did not see any strong correlations between the VP1 capsid proteins (data not shown). In contrast, many of the PyV TAgs were positively correlated, especially the phylogenetically related BK and JC, and HPyV6 and HPyV7 (Supplemental Figure 2.2). Due to the high correlation between TAgs, the immunodominance of VP1,[53] and the use of VP1 antibodies as cumulative exposure markers to PyVs, we focused our analyses on VP1 serostatus.

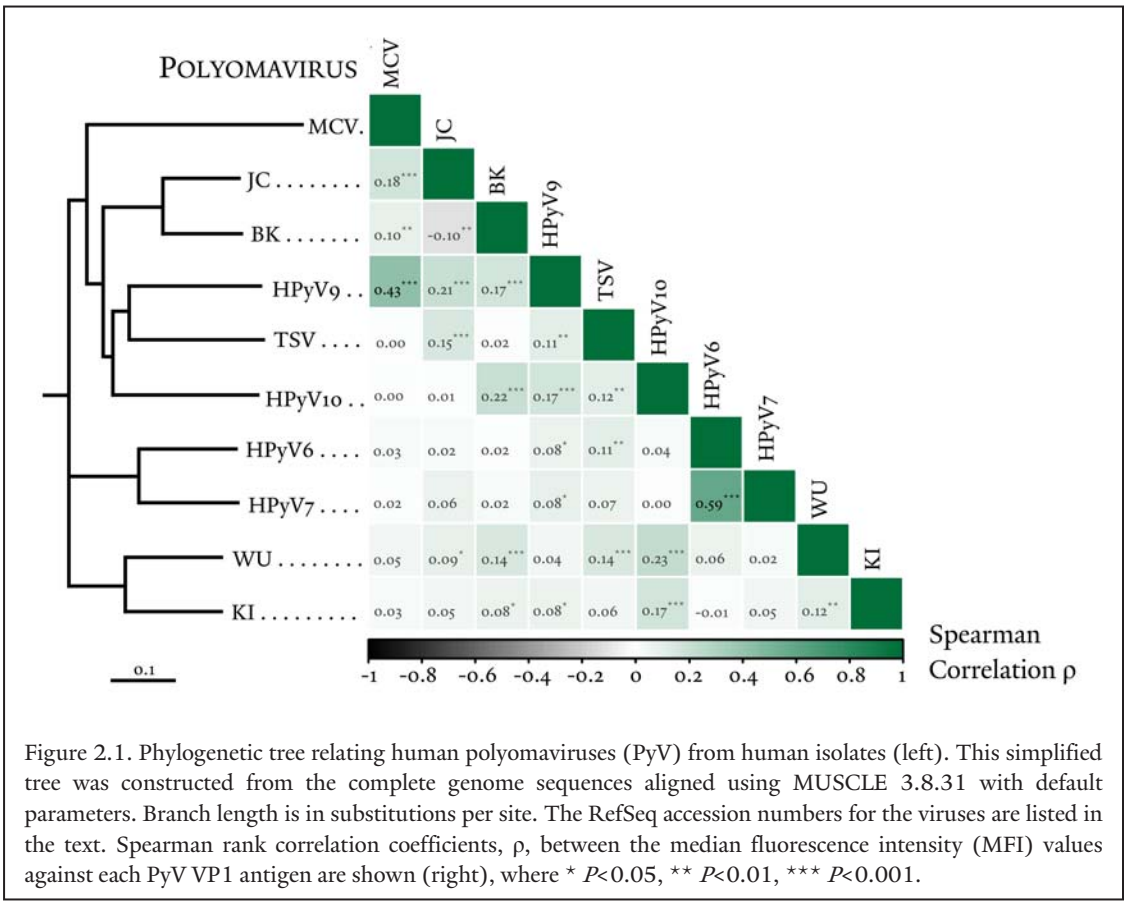



Figure 2.1. Phylogenetic tree relating human polyomaviruses (PyV) from human isolates (left). This simplified tree was constructed from the complete genome sequences aligned using MUSCLE 3.8.31 with default parameters. Branch length is in substitutions per site. The RefSeq accession numbers for the viruses are listed in the text. Spearman rank correlation coefficients, ρ, between the median fluorescence intensity (MFI) values against each PyV VP1 antigen are shown (right), where * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

All participants were seropositive for at least one type of PyV. Seroprevalences ranged from 17.6% for HPyV9 to 99.1 % for HPyV10 (Figure 2.2). The mean number of PyV VP1s to which individuals tested positive was 7.3 (standard deviation=1.4) (Figure 2.3), and increased with age group (*P*-for-trend=0.002). Men were seropositive for a higher average number of PyVs (mean=7.5, standard deviation=1.5) than women (mean=7.0, standard deviation=1.3; continuous *P*=0.0007, $X^2$ *P*=0.02, Table 2.1).

Figure 2.2. Seroprevalences of human polyomaviruses (PyV) as determined by VP1 seroreactivity among 460 study participants by age group. The overall seroprevalence (%) and number of participants who were seropositive for each PyV is shown above the bars. The number of participants within each age group is noted in the legend. * $P<0.05$, ** $P<0.01$, *** $P<0.005$ to test $P$-for-trend across age groups.

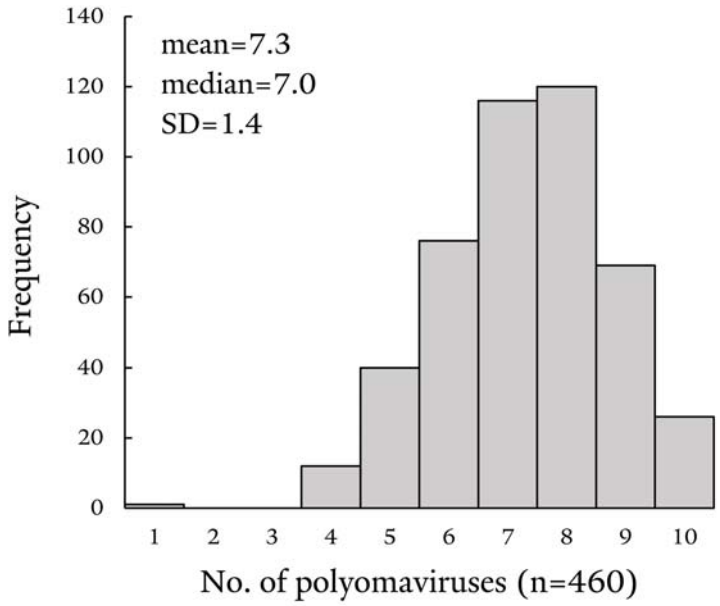



Figure 2.3. Number of different human polyomaviruses (PyV) against which an individual tested positive for the VP1 antigen amongst all 460 study participants. The mean number of different PyVs for which a participant was seropositive was 7.3 (standard deviation, SD=1.4) and median 7.0, out of 10 human PyVs assayed.

Seroprevalence was consistent across age groups for KI, WU, HPyV6, TSV, HPyV9, and HPyV10; and increased with age group for JC (*P*-for-trend=0.02), MCV (*P*=0.002), and HPyV7 (*P*=0.006) (Figure 2.2). A slight decrease by age group was observed for BK seroprevalence, albeit not of statistical significance (*P*=0.4). Men had higher VP1 seropositivity for many of the PyVs compared to women, and this difference was statistically significant for JC (60.0% in men and 48.9% in women, *P*=0.02), and TSV (85.4% in men and 73.9% in women, *P*=0.003) (Table 2.1).

Other individual characteristics were related to certain PyVs (Table 2.1). For HPyV9, a higher seroprevalence was observed among former or current smokers than never smokers (age group and gender-adjusted *P*=0.03). MCV seropositivity was less common among those who were current or former smokers, compared to those who had never smoked (adjusted *P*=0.02). HPyV7 seropositivity was more common among those who used glucocorticoids for one month

or longer (adjusted $P$=0.003). In addition, HPyV7 seropositivity was more common among those with skin that tended to tan following sun exposure, but this difference was not statistically significant (adjusted $P$=0.09). The sole organ transplant recipient was seropositive for almost all PyV types, except HPyV9. PyV seropositivity did not differ appreciably by level of education.

### 2.4.3 Polyomaviruses and immune cell proportions

Differences in immune cell proportions by serostatus were observed for BK, JC, and HPyV9 seropositivity (Figure 2.4). The median proportion of B cells was higher among those BK seropositive versus seronegative ($P$=0.03). In contrast, the median proportion of B cells was lower among those JC seropositive versus seronegative ($P$=0.02). CD8+ T cell proportions were lower for BK ($P$=0.02) and HPyV9 ($P$=0.05) seropositive versus those seronegative. However, the extent of these differences was relatively small. Correlations of the continuous MFI values for all PyV VP1 capsid and TAg proteins with immune cell proportions were small and not statistically significant (Supplemental Figure 2.2).

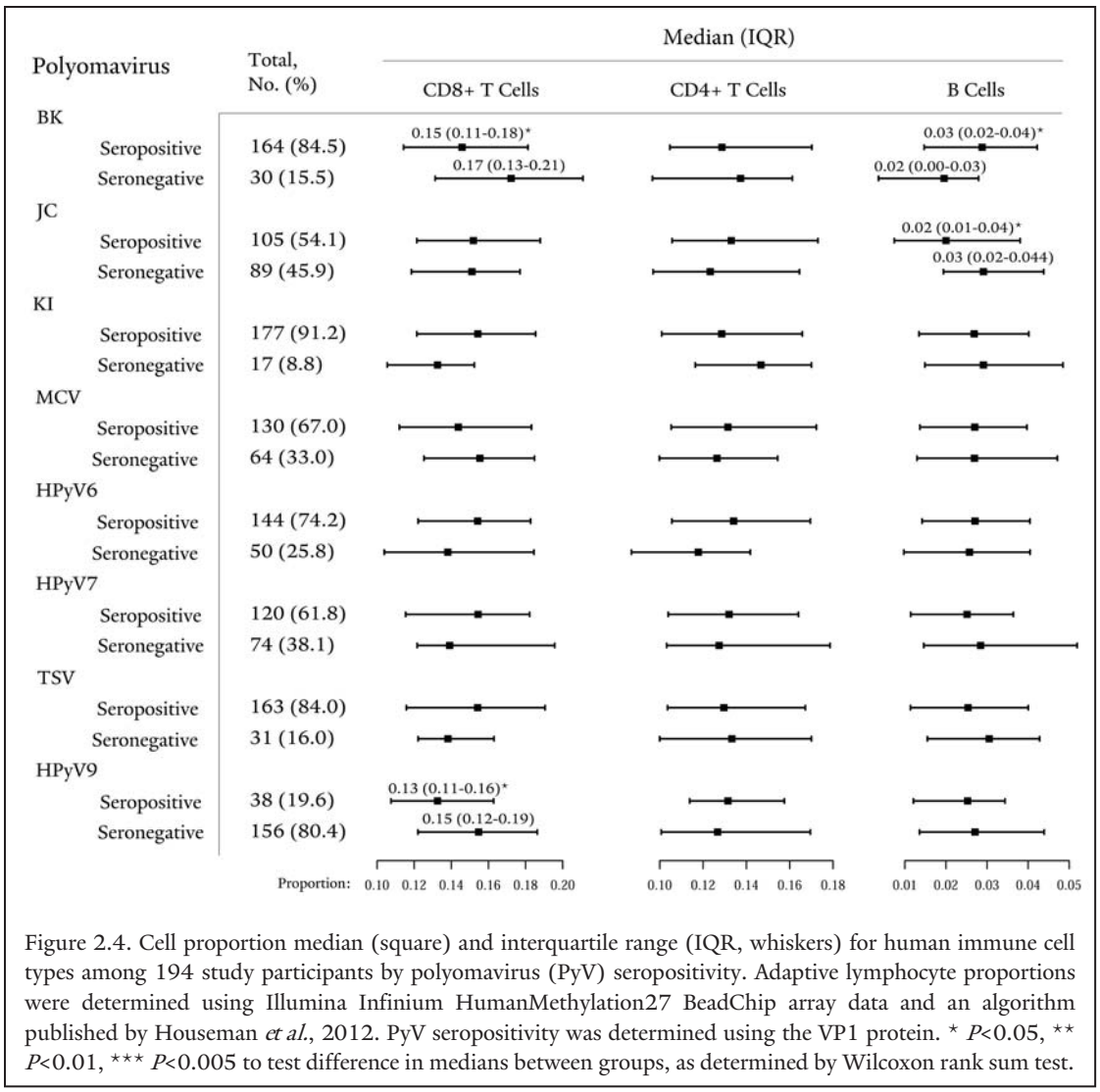



Figure 2.4. Cell proportion median (square) and interquartile range (IQR, whiskers) for human immune cell types among 194 study participants by polyomavirus (PyV) seropositivity. Adaptive lymphocyte proportions were determined using Illumina Infinium HumanMethylation27 BeadChip array data and an algorithm published by Houseman *et al.*, 2012. PyV seropositivity was determined using the VP1 protein. * $P<0.05$, ** $P<0.01$, *** $P<0.005$ to test difference in medians between groups, as determined by Wilcoxon rank sum test.

### 2.4.4 Polyomaviruses and DNA methylation

None of the locus-by-locus analyses reached statistical significance following FDR correction (Supplemental Figure 2.3). However, tens to thousands of genes were identified that contained a differentially methylated CpG locus without FDR correction ($P<0.05$), both for specific PyV types, and for number of PyV types seropositive above or below the median. In an exploratory gene function enrichment analysis, zinc finger proteins, transmembrane proteins (sialic acid binding), ion (potassium, calcium) channels, and cell death or apoptosis were among the major gene families consistently outputted by Gene Functional Classifications for all PyV types, and for

seropositivity above versus below the median number of PyV types. Moreover, immune-related gene families (such as immunoglobulin-like and interleukin receptors, immunoglobulin domains, immunoglobulin superfamily and interleukins) were returned for BK, JC, KI, MCV, HPyV6, HPyV9, and for number of PyVs above versus below the median; and pathways in cancer (such as oncogenes, RAS family members, carcinoembryonic antigens, and individual types of cancers) were returned for BK and HPyV9 (data not shown).

## 2.5 DISCUSSION

In a large sample of the general adult United States population, we examined the prevalence of antibodies against the VP1 capsid proteins from the first ten discovered human PyVs: BK, JC, KI, WU, MCV, HPyV6, HPyV7, TSV, HPyV9, and HPyV10. The number of PyVs to which individuals tested positive ranged from at least one to all ten, with a mean of 7.3 PyVs. Thus, our findings support the notion that PyVs are widespread in the human population. We also found age, gender, smoking status, glucocorticoid use, and indicators of immunity to be associated with PyV seropositivity.

We observed seroprevalences for BK, JC, WU, MCV, HPyV6, HPyV7, and TSV which were within the range of what has been reported previously in adults (reviewed in DeCaprio & Garcea, 2013,[4] Moens *et al.*, 2013,[43] and Nicol *et al.*, 2013 [48]). For KI (91.3%), the seroprevalence was just above the highest seroprevalences reported in the literature (55-90%),[3, 5, 14, 40, 41, 42, 43, 44] with the highest seroprevalence (90.0%) found in a United States female population from Seattle, Washington.[14] In Italian and German blood donors, the seroprevalence of HPyV9 ranged from 33-47%,[46, 48, 305] and was 24% in an Australian population,[282] whereas we found a seroprevalence of only 17.6%. For HPyV10, the seroprevalence was 42% in a large sample from Italy,[52] and 66% in a sample from Colorado, USA,[51] but almost 100% in our study.

However, a different HPyV10 was isolated from a human sample and investigated, with the study from Colorado using an isolate roughly 95% identical to the isolate used in the Italian study,[51, 52] and >99% identical to that used in our study,[33] which may impact antibody binding during serologic analysis. Thus, there may be variations in the PyV prevalences resulting from the use of different assays, as well as by geographical region and characteristics of the study population. Techniques used for PyV antibody detection and antigen preparation, and definition of cutoff values could also explain inter-study differences.

The trends we observed in seroprevalence by age were in agreement with the available literature. The increasing seroprevalence of JC and MCV with age was previously found in a large cohort from the Czech Republic[37] and among American blood donors.[306] The increase in HPyV7 seroprevalence with age was reported in an Italian study,[48] and may be explained by continuous transmission of HPyV7 throughout life.[48] The slight fall in BK seroprevalence we observed was present in English cohorts[36, 307] and a population survey in the Czech Republic,[37] and could be related to diminished antibody reactivity against BK over time or cohort effects in infection rates. For the PyVs that appeared stable during adulthood, primary infection likely occurred during childhood and thus little change in seroprevalence for adults might be expected. Indeed, there is evidence that primary infection occurs early in life for BK,[307] KI,[40, 42] WU,[40, 42] TSV,[41, 48] and HPyV10.[52]

Differences in PyV infection rates between men and women have generally not been observed in past studies.[5, 39, 42, 52, 307] However, an English study found more men than women to have antibodies against JC[307] as was found in our study. Differences in TSV seroprevalence by gender have not been reported previously[41, 48] and may be due to limited epidemiologic studies on TSV or population differences.

Studies examining the role of PyVs in smoking-related cancers have not found associations between smoking status and BK, JC, or MCV seropositivity across multiple populations.[308, 309, 310] We found seropositivity to MCV to be less common among smokers than never smokers – as observed in both a Chinese[310] and Chilean study[309] of non-small cell lung cancer samples assayed for MCV large TAg DNA, and in a Spanish study of bladder cancer and PyV seroprevalence;[308] although these differences were not statistically significant, possibly because these studies did not adjust for gender in their analyses.

We also found HPyV9 seropositivity to be higher among current and former smokers than never smokers, which has not been reported previously. In theory, cigarette smoking could increase risk of a viral infection through various mechanisms, including structural changes to the respiratory tract which may predispose to pathogen adherence; altered composition of cellular immunity that in turn could impair one's ability to limit viral replication (i.e., decreased number of CD4+ T cells and B cell proliferation, and eventually lower serum immunoglobulin levels); and enhanced cell infectivity (reviewed in Arcavi & Benowitz, 2004).[311] Interestingly, the prevalence of HPyV9 (first isolated from the blood and urine of a kidney transplant recipient[28]) appeared to increase following organ transplant – likely due to the immunosuppression.[312] As smoking may also be immunosuppressive, HPyV9 could infect or reactivate under conditions of immune dysfunction.

Indeed, in previous studies, the prevalence of multiple PyVs have differed in immunodeficient versus immunocompetent individuals. In our study, the single organ transplant recipient was seropositive for all but HPyV9: the least seroprevalent PyV).[28, 312] In numerous prospective studies, BK[313, 314] and JC[315] DNA or viral shedding have been detected in renal transplant patients at a higher frequency following transplantation, and antibody reactivity

against BK and JC was found to increase with time following transplant.[316] Additionally, BK replication has been associated with corticosteroid treatment in transplant recipients.[313, 317] We also found HPyV7 seropositivity to be more common among those who reported using glucocorticoids for at least one month, which have immunosuppressive properties.[318] Furthermore, a case series found high levels of JC viruria following long-term steroid use in an otherwise healthy adult;[319] however, additional data are needed. In a study on solid organ transplant recipients from Seattle, Washington, the reported seroprevalences for JC, KI, WU, MCV, HPyV6 and HPyV7 were lower than what we observed in our study population.[87] Thus, the effect of immunosuppression on PyV seropositivity will require additional study.

***Overall, methylation explained a relatively small proportion of variability in the pattern of PyV seroreactivity after adjusting for multiple comparisons in a subset of the study population. In our gene ontology (GO) analyses, we found altered DNA methylation in immune gene families related to PyV seropositivity. Changes in expression of immune-related genes was likewise observed in GO analyses of BK seropositivity in nephropathy patients from the Czech Republic.[320] In addition to genes related to carcinogenesis, GO analyses in our study returned possible associations between PyV seroreactivity and pathways involving sialic acids, zinc fingers, and apoptosis. Sialic acid binding by the VP1 subunit of the PyV structure assists with cell attachment for most PyVs,[321] the small TAg encodes two zinc fingers for protein binding,[322] and lytic infections caused by JC (leading to PML) and BK (leading to PVAN) involve tissue destruction.[7, 8]***

We further found individual PyVs to be associated with small but statistically significant changes in adaptive lymphocyte proportions: Those who were seropositive for BK had lower proportions of CD8+ T cells but higher B cell proportions, which was consistent with the findings that anti-allograft rejection immunosuppressive drugs often suppress T cell activation[323]

and are associated with PVAN.[324] HPyV9 seropositivity was also associated with lowered CD8+ T cell proportions. Furthermore, JC seropositivity was associated with lowered B cell proportions. Among immunocompromised (HIV/AIDS) patients, IgG responses against JC were higher among PML survivors than those who died,[325] suggesting B cells and the humoral immune response may be important in defense against JC. Impaired cellular immune responses and decreased immune surveillance could promote active PyV infection, as has been known for BK and JC.[318] Smaller differences in immune cell proportions were observed in our study, and this might be expected among immunocompetent individuals. Thus, while our findings are speculative, they are consistent with the clinical and experimental evidence[4] suggesting that PyV infectivity may be influenced by conditions of immunosuppression and warrant further investigation as potential oncogenic viruses.

### 2.5.1 Strengths and limitations of the study

An advantage of our study was the availability of a large number of samples from participants in a population-based study. The comprehensive assessment of multiple human PyV types in a North American study population, and the association of individual PyV seropositivity with sociodemographic characteristics as well as measures of immune cell proportions derived from a novel algorithmic technique, were unique strengths of our research.

Limitations of our research include the use of PyV seroreactivity as a measure of PyV infection.[39, 59] However, the GST capture of recombinantly expressed VP1 capsid proteins was shown to be a reliable technique to assess PyV seroreactivity and has been used as a marker of PyV infection in prior studies.[39, 53, 54, 59] In addition, MCV viral load and antibody titer have been shown to have a strong positive monotonic correlation,[49] although less is known about the other PyV types. The VP1 assay showed minimal signs of cross-reactivity in antibody detection

between PyV types, although potential serological cross-reactivity between HPyV6 and HPyV7, and MCV and HPyV9, may limit our interpretation for these PyV types. However, once we restricted to those where were just PyV seropositive, we did not see any VP1 cross-reactivity. Although it is possible that our seroprevalences could be affected by our choice of cut point to convert a continuous MFI measure into a binary variable, our sensitivity analysis suggested that this did not appreciably impact our findings. Cross-reactivity with other, yet undiscovered, human PyVs cannot be excluded. For instance, we did not investigate the recently identified Saint Louis polyomavirus (STLPyV),[326] HPyV12,[327] or New Jersey polyomavirus (NJPyV).[328] Finally, the immune cell proportions were derived from genome-wide DNA methylation profiles[293] rather than more labor intensive methods such as flow cytometry, although cell lineage-specific DNA methylation patterns closely approximate leukocyte subsets purified from human blood.[329] As the same blood sample was used both for PyV antibody and DNA methylation assays, we cannot determine whether immune cell variability is a cause or consequence of seroprevalence.

### 2.5.2 Conclusions

Our findings suggest that PyVs may be ubiquitous in the United States population. The seroprevalence of specific PyV types may differ by individual characteristics (such as age, gender, smoking status, glucocorticoid use), and immune profiles, which could impact disease risk.

# Chapter 3

# Human polyomaviruses and incidence of cutaneous squamous cell carcinoma in the New Hampshire Skin Cancer Study

*Published as: Anala Gossai, Tim Waterboer, Anne G. Hoen, Shohreh F. Farzan, Heather H. Nelson, Angelika Michel, Martina Willhauck-Fleckenstein, Brock C. Christensen, Ann E. Perry, Michael Pawlita, and Margaret R. Karagas. "Human polyomaviruses and incidence of cutaneous squamous cell carcinoma in the New Hampshire Skin Cancer Study." Cancer Medicine (2016): 1-12. (excluding sections 3.3.3, 3.3.5, 3.3.6, 3.4.3, 3.4.4, & **bold italicized** paragraphs of 3.5).*

## 3.1 ABSTRACT

Squamous cell carcinoma (SCC) of the skin is a malignancy arising from epithelial keratinocytes. Experimental and epidemiologic evidence raise the possibility that human polyomaviruses (PyV) may be associated with the occurrence of SCC. To investigate whether the risk for SCC was associated with PyV infection, seropositivity to ten PyV types was assessed following diagnosis in a population-based case-control study conducted in the United States. A total of 253 SCC cases and 460 age group and gender-matched controls were included. Antibody response against each PyV was measured using a multiplex serology-based glutathione S-transferase capture assay of recombinantly expressed VP1 capsid proteins. Odds ratios (OR) for SCC associated with seropositivity to each PyV type were estimated using logistic regression, with adjustment for potentially confounding factors. SCC cases were seropositive for a greater number of PyVs than controls ($P$=0.049). Those who were JC seropositive had increased odds of SCC when compared to those who were JC seronegative (OR=1.37, 95% CI: 0.98-1.90), with an increasing trend in SCC risk with increasing quartiles of seroreactivity ($P$-for-trend=0.04). Furthermore, a multifactor dimensionality reduction (MDR) analysis suggested effects on the risk of SCC from synergistic interactions between JC, MCV, and age group (OR=1.7, 95% CI: 1.2-2.4). There were no clear associations between SCC risk and serostatus for other PyV types. This study provides limited evidence that infection with certain PyVs may be related to the occurrence of SCC in the general population of the USA.

## 3.2 INTRODUCTION

Cutaneous squamous cell carcinoma (SCC) is a keratinocyte cancer (KC), with increasing incidence rates reported in the USA,[101, 102, 103, 104] Australia,[105, 106] and Europe.[107, 108] Estimates of up to 420,000 individuals in the USA were diagnosed with incident SCC in 2012, and 8,800 SCC patients died from this malignancy.[102] Established risk factors include elderly age[101, 104, 330] and ultraviolet light,[96, 170] but environmental exposure to ionizing radiation,[173, 174] arsenic,[175, 176, 177, 178] and polycyclic aromatic hydrocarbons[179, 180] further increases the risk of SCC. Inherited traits such as skin pigmentation[171, 172] and genetic defects in the ability to correct DNA photolesions (e.g., xeroderma pigmentosum)[211, 212, 331] also contribute to risk. While the prognosis is generally favorable, metastases can occur; and the aging population, as well as disfigurement and morbidity resulting from SCC, makes this malignancy an important public health concern and under-recognized health burden.

Immunocompromised persons (e.g., organ transplant recipients, and those given immunosuppressant medications) carry a higher risk of SCC,[84, 101, 104, 213, 214, 332] raising the possibility of a viral etiology. However, the detection of viral sequences in tumors can reflect either a causal or bystander role.[333, 334] Thus far, genus β human papillomaviruses (HPV)[61, 278] have been associated with an increased risk of SCC, but a causal relationship has not yet been established in the general population.[335]

A potential etiologic role for polyomaviruses (PyV) in cancer has been investigated, and a rapidly expanding number of viral types are being identified in the family *Polyomaviridae* (reviewed in DeCaprio & Garcea, 2013).[4] PyVs are DNA viruses with an icosahedral capsid ~45 nm in diameter containing a circular double-stranded genome[2, 3] that encodes capsid proteins (VP1, VP2, VP3), as well as small and large T antigens (TAg).[2] In simian virus 40 (SV40) – a

PyV naturally infecting Asian macaques,[336, 337] the large TAg possesses tumorigenic properties, including the ability to bind and thereby inactivate tumor suppressor proteins Rb[15] and p53,[16] thus stimulating the host cell cycle.[2]

While the ability of PyVs to cause tumors *in vitro* and in experimental systems is undisputed, their role in human malignancies – and specifically KCs – is just beginning to emerge. Multiple human PyVs show evidence of skin tropism, including Merkel cell polyomavirus (MCV), *Trichodysplasia spinulosa*-associated polyomavirus (TSV), and human polyomaviruses 6 and 7 (HPyV6 and HPyV7).[3] MCV was initially discovered in Merkel cell carcinoma (MCC) of the skin,[24] and integration of the viral genome in MCC tumors[338] with potentially carcinogenic mutations in the large TAg[12] supported a causal role for MCV in MCC development. Recently, a clinic-based case-control study from Florida, USA, found an increased SCC risk associated with antibodies against MCV measured following diagnosis, and the association was stronger among those with tumors containing MCV viral sequences.[54] Conversely, a case-control study conducted within organ transplant recipients found no evidence of an association between SCC development following transplant surgery and the seroprevalence of numerous PyV types prior to transplantation.[87] Further, a large prospective Swedish study found no association between SCC and JC seropositivity assessed prior to diagnosis.[339]

Limited epidemiologic research exists on the role of PyV infections (apart from MCV) in SCC carcinogenesis. We therefore sought to investigate the potential association between PyVs and SCC by performing a comprehensive serologic analysis of the frequency of antibodies to the first ten identified human PyVs: BK virus;[21] JC virus;[20] Karolinska Institute polyomavirus (KI);[22] Washington University polyomavirus (WU);[23] MCV;[24] HPyV6 and HPyV7;[27] TSV;[9] human polyomavirus 9 (HPyV9);[28] and Malawi/human polyomavirus 10 (HPyV10)[30] in

plasma samples collected as part of a large, population-based case-control study conducted in New Hampshire, USA.

## 3.3 METHODS

### 3.3.1 Study population

The New Hampshire Skin Cancer Study population and methods have been described in detail elsewhere (epidemiologic study design described in Karagas *et al.*, 1998 [340]).[62, 120, 278] Briefly, histologically-confirmed, incident SCC cases were identified through active surveillance of dermatology and pathology laboratories throughout New Hampshire, USA. We selected all identified SCC cases diagnosed between July 1993 and June 1995, during the first enrollment phase of this large, population-based case-control study. Controls were selected from lists of New Hampshire residents obtained from the New Hampshire Department of Transportation (<65 years) and Medicare enrollment lists (≥65 years), and frequency-matched to the age (25-34, 35-44, 45-54, 55-64, 65-69, and 70-74 years) and gender distribution of cases. For the purpose of the interview, controls were randomly assigned a reference date that matched a case's diagnosis date. To be eligible, participants were required to be residents of New Hampshire, aged 25-74 years at time of diagnosis, speak English, and have a listed telephone number. We excluded participants with squamous cell or basal cell carcinomas on genital sites.

Study participants completed an extensive, structured interview, usually in their homes. Personal interviews were conducted to obtain sociodemographic information (e.g., level of education), lifestyle factors (e.g., cigarette smoking), sunlight-related characteristics (e.g., response to first exposure in summer to one hour of sunlight, number of severe sunburns, and skin color), and medical history (e.g., prior history of skin cancer, use of oral glucocorticoids for one month or longer, and organ transplant status). Data on the primary tumor(s) (e.g., SCC

anatomical location) were collected from a medical records review. Tumors diagnosed as a recurrence of a previously treated tumor, and those appearing contiguous with a scar from a previously excised skin cancer, were not considered a new primary SCC and thus were excluded. All participants provided informed consent in accordance with the Committee for the Protection of Human Subjects at Dartmouth College.

### 3.3.2 Human polyomavirus serology

As part of the New Hampshire Skin Cancer Study, we requested a blood sample from all participants for use in future research.[340] We collected venous blood samples of 20-30 mL in heparinized tubes following SCC diagnosis (as described in Karagas *et al.*, 2010 and Karagas *et al.*, 2006 [62, 278]). Blood was separated by centrifugation at 2,500*g* for 20 minutes at 4°C, and each component (plasma, red blood cells, buffy coat) was labeled and stored separately at -80°C until analysis. Specimen label did not reveal the case-control status of the study participant. Samples were shipped to the German Cancer Research Center (DKFZ; Heidelberg, Germany) on dry ice for analysis.

Plasma samples were assayed for antibodies against the immunodominant VP1 capsid protein[53] of ten human PyVs (BK, JC, KI, WU, MCV isolate 344, HPyV6, HPyV7, TSV, HPyV9, HPyV10). Plasma samples were also tested for antibodies against the TAg of selected PyV types (large TAg for BK, JC, MCV, HPyV6, HPyV7, TSV, HPyV10; and small TAg for MCV). In our prior study, we did not find strong positive correlations or evidence of cross-reactivity between the VP1 capsid proteins of most PyV types (data not shown), suggesting that risk estimates obtained during analysis would be specific to that PyV type.[341] However, the strong positive correlations between TAg seroreactivities from various PyVs suggestive of assay cross-reactivity (Supplemental Figure 3.1), and the small number of participants TAg

seropositive (Supplemental Table 3.1), resulted in the exclusion of TAg serostatus from the presented analyses. The multiplex antibody detection approach was based on a glutathione S-transferase (GST) capture enzyme-linked immunosorbent assay (ELISA) method in combination with fluorescent bead technology (Luminex Corp., Austin, Texas, USA).[55, 58] Antigen preparation and techniques used for PyVs[39, 53, 54] closely follow methods applied to HPVs as described previously.[55, 56]

Seroreactivity against PyV VP1 proteins was expressed as the median fluorescence intensity (MFI) of 100+ beads of the same internal color.[58] MFI values reflect antibody affinity, titer, and reactivity determined by dilution series.[60] Standard cut points to define seropositivity were chosen for each PyV by visual inspection of frequency distribution curves (percentile plots) for the inflection points of all sera tested, as done in prior studies.[14, 39, 56, 62] The standard cutoff value for VP1 was 250 MFI units for all ten PyVs (as used in Teras *et al.*, 2014 [281] and Gossai *et al.*, 2015 [341]). To evaluate the robustness of odds ratio (OR) estimates for SCC by PyV seropositivity, we used a sliding cut point between 50 and 450 MFI units, and also calculated cut points from controls using a method adapted from van der Meijden *et al.*, 2013 [282] (Supplemental Figure 3.2). Given the stability of ORs to cut point definition, we ultimately used the standard cut points in all analyses.

### 3.3.3 *HLA-C* and *NKG2D* genotyping

As the immune system and natural killer (NK) cell killing phenotype is related to both risk of SCC and potentially PyV infection outcomes, we tested the hypothesis that genetic variation in *HLA-C* and *NKG2D* is related to PyV seropositivity amongst controls, and modifies the risk of SCC by PyV seroreactivity in a population-based case-control study. *HLA-C* encodes a ligand for inhibitory killer immunoglobulin-like receptors (KIR).[342] γδ T cells express activating NK group

2 member D (NKG2D) receptors, which respond to stress proteins.[342] The *NKG2D* genotype is associated with *NKG2D* expression levels and NK cell cytotoxicity.[342]

DNA was extracted from the buffy coat of blood samples for multiplex immune single nucleotide polymorphism (SNP) genotyping for *HLA-C* and *NKG2D* with a Luminex-based typing kit (Lifecodes) and a Taqman assay (Life Technologies), as has been described previously.[343, 344, 345] Briefly, the inhibitory ligand for *KIR* receptors, *HLA-C* alleles, were classified as group C1 when containing serine (AGC) at codon 77 and asparagine (AAC) at codon 80, and group C2 when containing asparagine (AAC) at codon 77 and lysine (AAA) at codon 80, in the ligands. *HLA-C -35 kb* is the *HLA-C* expression SNP (CC is the high expressing genotype and TT has low activity). *NKG2Da* corresponds to rs2255336 (C=low activity, T=high activity, with CC having the lowest activity), and *NKG2Db* corresponds to rs1049174 (C=low activity, G=high activity, with CC having the lowest activity).

### 3.3.4 Statistical analyses

Individual characteristics of SCC cases and controls were compared using the $X^2$ test (for categorical variables, i.e., gender, education, smoking status, skin color, skin sun sensitivity, number of sunburns, prior KC, glucocorticoid use) or Fisher's exact test (for categorical variables with small strata, i.e., transplant recipients), and Wilcoxon rank sum test (for continuous variables, i.e., age, mean number of PyVs seropositive). Among controls, we previously published the seroprevalence of each PyV type, and tested the association between various individual characteristics in relation to PyV seropositivity; therefore, these analyses were not repeated in this study.[341]

We used unconditional logistic regression to calculate the ORs and 95% confidence intervals (CI) for SCC by VP1 seropositivity compared to seronegativity for each PyV type, while

adjusting for age group and gender (as used in the frequency matching). We also examined seropositivity for multiple PyV types (7-8 and 9-10 types positive compared with 1-6 types positive), and calculated a *P*-for-trend based on these categories and a continuous variable of the number of types seropositive. Quartiles of seroreactivity based on the control distributions of continuous MFI values were created for each PyV, and associated with SCC by comparing the second, third, and fourth quartile to the first (lowest) quartile. Tests-for-trend were conducted by including an ordinal variable in the logistic model. In addition, we evaluated seropositivity for multiple cutaneous PyVs (i.e., MCV, HPyV6, HPyV7, TSV), and whether seropositivity for all 4 or ≤3 cutaneous PyVs, was associated with SCC risk.

Potentially confounding covariates included level of education, smoking status, eye color, hair color, skin color, skin sensitivity to the sun, number of lifetime painful sunburns, lifetime sun exposure, and the self-reported use of oral glucocorticoids for one month or longer. As none of these factors consistently produced a >10% change in OR[284] when individually placed in a logistic regression model including the frequency matching factors (age group and gender), our OR estimates were ultimately adjusted for only these factors.

We assessed the potential modifying effects of prolonged oral glucocorticoid use for reasons other than organ transplantation, as PyVs have been related to skin conditions and cancers in immunosuppressed populations. In these stratified analyses, we classified participants as users if they reported taking glucocorticoids for one month or longer, and excluded those who reported having had an organ transplant. An association between SCC risk and seropositivity to multiple cutaneous β HPVs has been reported previously.[61, 278] Therefore, using published serologic data from the New Hampshire Skin Cancer Study[61, 278] on 16 cutaneous β HPVs (specifically 5, 8, 20, 24, 36 (beta-1); 9, 15, 17, 23, 38, 107 (beta-2); 49, 75, 76 (beta-3); 92

(beta-4); and 96 (beta-5)), we stratified participants by those with seropositivity to 0-1 or seropositivity to ≥2 β HPV types to assess possible effect modification by concomitant β HPV infections. We further calculated separate ORs for SCC occurring exclusively on anatomical sites with chronic sunlight exposure (head or neck) and other body, sites in comparison to controls.

As sensitivity analyses, we excluded participants with a history of organ transplantation (n=7), restricted to participants with no previous skin cancers (neither a prior SCC nor basal cell carcinoma (BCC); n=423 controls and 179 cases), and excluded participants with a concomitant BCC (n=22 cases) to assess whether the OR estimates differed from those obtained for all participants. No appreciable change in results was detected (Supplemental Figure 3.3), and thus these individuals were included in the presented analyses. All statistical tests were two-sided, and significance was assessed at the $\alpha=0.05$ level. Statistical analyses were performed in R version 3.0.1.

### 3.3.5 Multifactor dimensionality reduction analysis

We systematically examined the effect of multiple interactions between PyVs on the risk of SCC using the machine learning approach, Multifactor Dimensionality Reduction (MDR version 3.0.3_dev, www.epistasis.org). MDR was developed as a model-free approach to detect epistasis through the reduction of high-dimensional data by assigning all parameter combinations to either "high risk" or "low risk" groups for having the outcome of interest (based on the ratio of cases to controls).[346, 347, 348] Here, MDR was used to examine SCC case or control status by seropositivity for 10 human PyVs. MDR assigned PyV combinations to "high risk" or "low risk" groups for having SCC.[346, 347, 348] Within MDR, the simple method of duplicating rows was used to adjust for age group and gender.[349] The type of relationships between PyVs was then inferred using an entropy-based approach, which computed information gained from considering interactions

between parameters: a negative entropy value implied redundancy (that is, the effect of the two interacting parameters is smaller than a sum of individual effects), while a positive entropy value implied synergy (that is, the effect of the two interacting parameters is greater than a sum of individual effects).[348]

All ≤5-way PyV combinations were investigated using an exhaustive search. PyV combinations, as well as age group and gender, were investigated using an exhaustive interaction method (up to 5-way variable combinations), and models were compared to avoid overfitting or overparameterization, and to increase model generalizability. After training on 9/10ths of the dataset, the best model was required to have the highest balanced testing accuracy (TA) in classification of a SCC case or control, a cross validation (CV) consistency between 9/10-10/10 (using a 10-fold CV), and remain statistically significant after undergoing 1000-fold permutation testing (i.e., the case-control labels were shuffled 1000 times and the entire MDR model was fitted on each randomized dataset). Variables included in the "best model" were then assessed in logistic regression models.

### 3.3.6 Immuno-genotype statistical analyses

We first solved the Hardy-Weinberg equation for each allele within control and case participants, separately. We then tested the association between various individual characteristics (including case or control status) in relation to genotype amongst all study participants using the $X^2$ test (for categorical variables), or Fisher's exact test (for categorical variables with small strata, i.e., <10 participants), and Wilcoxon rank sum or Kruskal-Wallis tests (for continuous variables). Amongst controls, we examined the frequency of PyV seropositivity for each type by genotype using the $X^2$ test or Fisher's exact test (for small strata). We also used the Kruskal-Wallis test to investigate the association between genotype and number of PyV types to which an individual

tested positive. The association between PyV serostatus (outcome) and genotype (predictor) amongst controls was further assessed using unconditional logistic regression to adjust for the potentially confounding influence of age group and gender.

We stratified participants by *HLA-C* and *NKG2D* genotypes to look for modifying effects of inherent immunity on the relationship between PyVs and SCC. Stratified by immuno-genotype, we used unconditional logistic models to calculate the odds of 95% confidence intervals for SCC by VP seropositivity compared to seronegativity for each PyV type, while adjusting for age group and gender (the matching factors) to assess for potential effect modification within each genotype. In addition, quartiles of VP1 seroreactivity based on the control distributions of continuous MFI values were created for each PyV, and associated with SCC by comparing the second, third, and fourth quartile to the first (lowest) quartile, for each genotype, with adjustment for the matching factors. All statistical tests were two-sided, with significance assessed at the $\alpha$=0.05 level, and analyses were performed in R version 3.1.0.

## 3.4 RESULTS

### 3.4.1 Study population characteristics

Plasma samples for PyV serology were obtained from 253 (86.3%) of the 293 interviewed SCC cases following diagnosis, and 460 (85.2%) of the 540 interviewed controls (excluding 1 control with an insufficient bead count during serologic analysis). No appreciable differences were noted in the characteristics of individuals for whom we did not obtain serology data (data not shown). Compared to controls, SCC cases were older, were of lighter skin color, had skin that tended to burn following sun exposure, reported a greater number of painful sunburns in their lifetime, were more likely to have had a prior KC (SCC and/or BCC), and were more likely to have had an organ transplant (Table 3.1).

Table 3.1. Selected characteristics of cutaneous squamous cell carcinoma (SCC) cases and controls from the New Hampshire Skin Cancer Study (n=713).[a]

| Variable | SCC Cases (n=253), No. (%) | Controls (n=460), No. (%) |
|---|---|---|
| Gender | | |
| Male | 168 (66.4) | 280 (60.9) |
| Female | 85 (33.6) | 180 (39.1) |
| Median age, SD (years) | 68 (8.0) | 65 (10.7)*** |
| Education | | |
| Elementary to high or technical school | 111 (43.9) | 227 (49.3) |
| Any college | 80 (31.6) | 144 (31.1) |
| Graduate or professional school | 61 (24.1) | 89 (19.3) |
| Smoking status [b] | | |
| Never | 79 (31.2) | 146 (31.7) |
| Former | 133 (52.6) | 230 (50.0) |
| Current | 40 (15.8) | 84 (18.3) |
| Skin color | | |
| Light | 213 (84.2) | 279 (60.6)*** |
| Medium | 38 (15.0) | 180 (39.1) |
| Skin sun sensitivity [c] | | |
| Severe sunburn with blistering | 22 (8.7) | 28 (6.1)*** |
| Painful sunburn and then peeling | 93 (36.8) | 116 (25.2) |
| Mild sunburn with some tanning | 111 (43.9) | 234 (50.9) |
| Tan without sunburn | 25 (9.9) | 80 (17.4) |
| No. of lifetime painful sunburns [d] | | |
| 0 | 66 (26.1) | 147 (32.0)*** |
| 1-2 | 49 (19.4) | 134 (29.1) |
| 3+ | 136 (53.8) | 174 (37.8) |
| Prior keratinocyte cancer [e] | | |
| Yes | 74 (29.2) | 37 (8.0)*** |
| No | 179 (70.8) | 423 (91.9) |
| Glucocorticoid use [f] | | |
| Yes | 33 (13.0) | 39 (8.5) |
| No | 211 (83.4) | 415 (90.2) |
| Transplant recipient | | |
| Yes | 6 (2.4) | 1 (0.2)** |
| No | 246 (97.2) | 458 (88.6) |

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. $P$ values obtained from $X^2$, Fisher's exact, or Wilcoxon rank sum test (as appropriate) comparing sociodemographic and skin cancer risk factors between SCC cases and controls.

[a] Numbers may not sum to the overall total due to missing data. They were excluded from complete-case analyses.

[b] Cigarette smoking status at 1 year prior to the reference or diagnosis date

[c] Sun sensitivity was defined as the skin reaction to 1 hour of sun exposure the first time in the summer.

[d] Sunburns that caused pain for 2 or more days

[e] Prior keratinocyte cancer was defined as having had a previous squamous cell, basal cell, or both squamous cell and basal cell carcinoma.

[f] Glucocorticoid use was defined as having used oral steroid or corticosteroid medications (e.g., cortisone or prednisone) for one month or longer.

### 3.4.2 Polyomaviruses and cutaneous squamous cell carcinoma

Seroprevalences ranged from 17.6% for HPyV9 to 99.1% for HPyV10 among controls, and from 22.9% for HPyV9 to 98.8% for HPyV10 among cases (Figure 3.1). All study participants were seropositive for at least one type of PyV, but cases (mean=7.52, standard deviation=1.40) were seropositive for a slightly greater number of PyVs than controls (mean=7.30, standard deviation=1.45; Wilcoxon rank sum test *P*=0.049). However, we found no evidence of an increasing trend in SCC odds with increasing number of PyV types for which an individual tested seropositive, either using a categorical variable (*P*-for-trend=0.42) or on a continuous scale (*P*=0.32) (Table 3.2). Further, no association with SCC was observed for combined seropositivity to the known cutaneous PyVs (i.e., MCV, HPyV6, HPyV7, and TSV; Supplemental Table 3.3).

Table 3.2. Odds ratios (95% confidence intervals) for cutaneous squamous cell carcinoma (SCC) by number of polyomavirus (PyV) types seropositive among 713 study participants from the New Hampshire Skin Cancer Study.

| No. of PyV types seropositive | Controls (n=460), No. (%) | SCC Cases (n=253) | |
|---|---|---|---|
| | | No. (%) | OR (95% CI)* |
| 1-6 | 129 (28.0) | 60 (23.7) | 1.00 (referent) |
| 7-8 | 236 (51.3) | 128 (50.6) | 1.04 (0.71-1.53) |
| 9-10 | 95 (20.6) | 65 (25.7) | 1.21 (0.76-1.91) |
| *P* for trend | | | 0.42 |
| Continuous [mean No. (SD)] | 7.30 (1.45) | 7.52 (1.40) | 1.06 (0.95-1.18) |
| *P* | | | 0.32 |

* Adjusted for age group and gender. OR=odds ratio, CI=confidence interval, SD=standard deviation.

In an analysis of SCC associations with antibodies to specific capsid antigens from each PyV type (classified as a dichotomous variable), an elevated odds of SCC was observed with seropositivity to JC of borderline statistical significance (OR=1.37, 95% CI: 0.98-1.90; Figure 3.1), and a positive trend in SCC risk was associated with increasing MFI quartiles of JC

seroreactivity (*P*-for-trend=0.04) (Table 3.3). A weakly positive association with SCC risk was observed for seropositivity to WU (OR=2.04, 95% CI: 0.64-6.49) (Figure 3.1), and a slightly inverse association with SCC risk was observed for seropositivity to HPyV10 (OR=0.72, 95% CI: 0.16-3.32), but neither estimates achieved statistical significance. There were no trends in SCC risk by MFI quartiles for any PyV type but JC (Supplemental Table 3.2).

| Polyomavirus: | | BK | JC | KI | WU | MCV | HPyV6 | HPyV7 | TSV | HPyV9 | HPyV10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Seropositive No. (%) | Controls: (n=460) | 403 (87.6) | 256 (55.6) | 420 (91.3) | 448 (97.4) | 324 (70.4) | 339 (73.7) | 259 (56.3) | 372 (80.9) | 81 (17.6) | 456 (99.1) |
| | SCC Cases: (n=253) | 217 (85.8) | 167 (66.0) | 237 (93.7) | 249 (98.4) | 174 (68.8) | 190 (75.1) | 156 (61.7) | 204 (80.6) | 58 (22.9) | 250 (98.8) |

* Adjusted for age group and gender. CI=confidence interval.

Figure 3.1. Odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among 713 study participants from the New Hampshire Skin Cancer Study. The overall seroprevalences (%) and number of participants who were seropositive for each PyV is shown below the plot area separately for cases and controls.

Table 3.3. Odds ratios (95% confidence intervals) for cutaneous squamous cell carcinoma (SCC) by quartiles of JC polyomavirus (PyV) seroreactivity among 713 study participants from the New Hampshire Skin Cancer Study.

| PyV seroreactivity (MFI units) | Controls (n=460), No. (%) | SCC Cases (n=253) | |
|---|---|---|---|
| | | No. (%) | OR (95% CI)* |
| **JC** | | | |
| Quartile 1 | 115 (25.0) | 51 (20.2) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 43 (17.0) | 0.79 (0.48-1.30) |
| Quartile 3 | 115 (25.0) | 74 (29.2) | 1.27 (0.80-1.99) |
| Quartile 4 | 115 (25.0) | 85 (33.6) | 1.41 (0.90-2.20) |
| *P* for trend | | | 0.04 |

* Adjusted for age group and gender. OR=odds ratio, CI=confidence interval.

SCC was more strongly related with seropositivity to JC among those reporting oral glucocorticoid use for one month or longer (OR=1.88, 95% CI: 0.61-5.76) than those without a history of prolonged glucocorticoid use (OR=1.37, 95% CI: 0.96-1.96) – albeit with limited statistical power (*P*-for-interaction=0.50; Supplemental Figure 3.4). When excluding glucocorticoid users, the positive association between SCC risk and HPyV9 seropositivity was strengthened (OR=1.49, 95% CI: 0.99-2.24). Interestingly, a weakly positive association with SCC risk was found for HPyV10 seropositivity among those seropositive for <2 β HPV types (OR=1.45, 95% CI: 0.15-14.37), although with wide confidence intervals. We did not find any notable differences for the other PyVs according to glucocorticoid use, anatomical site of the SCC tumor, or seropositivity to β HPVs, in stratified analyses.

### 3.4.3 MDR findings

MDR analysis for SCC consistently revealed the best model to include 3 variables: JC, MCV, and age group. The model had a balanced accuracy of 0.6 (slightly better than random), CV consistency of 10/10, and *P*=0.007-0.008 after 1,000 permutations. From the graphical model, it was inferred that a synergistic interaction between JC and age group, and this interaction with

MCV, was detected (Figure 3.2A). There were 302 participants allocated to the high risk group, and 578 participants allocated to the low risk group in MDR (the sum of MDR participants was slightly higher than the actual sample size due to the row duplication procedure used for covariate adjustment; Figure 3.2B).[349] The risk of SCC amongst those who were in the high risk group was OR=2.5 (95% CI: 1.9-3.4) times greater than the risk of SCC amongst those who were in the low risk group.

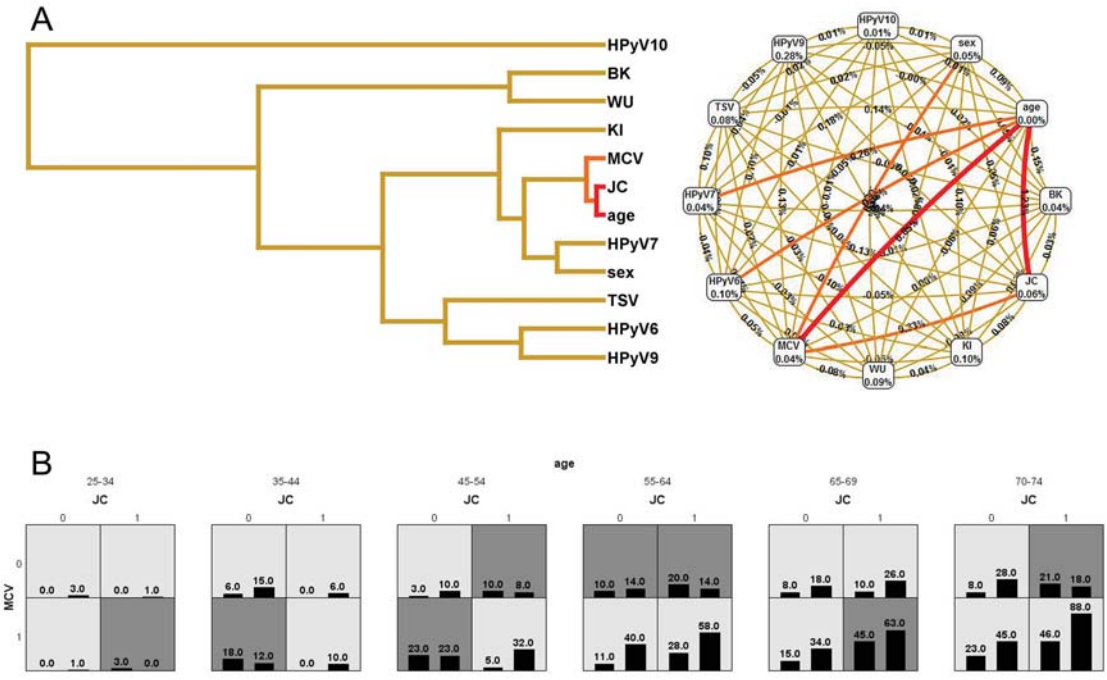



Figure 3.2. Top Multifactor Dimensionality Reduction (MDR) result for cutaneous squamous cell carcinoma (SCC) case/control status by polyomavirus (PyV) serostatus for 10 human PyV types, as well as age group and gender. The overall model balanced accuracy=0.6. The risk of SCC amongst those who were in the high risk group was OR=1.7 (95% CI: 1.2-2.4) times greater than the risk of SCC amongst those who were in the low risk group, following adjustment for just gender amongst the n=713 study participants. **A.** Entropy dendrogram (left) and circle graph (right). Red and orange lines suggest a synergistic relationship, while yellow lines suggest independence. For the dendrogram, the shorter the line connecting two attributes, the stronger the interaction. Information gained or lost is shown in the circle graph. **B.** Graphical model. High-risk cells are shaded dark gray while low-risk cells are shaded light gray, with the number of cases (left bars) and controls (right bars) for each cell shown. JC and MCV serostatus is noted as seropositive (1) or seronegative (0). Age group (years) is noted along the top.

When the classification rules were used to assign study participants to risk groups in the original dataset, 107 cases and 138 controls were allocated to the high risk group, and 146 cases

and 322 controls were allocated to the low risk group (n=713), resulting in the risk of SCC amongst those at high risk being OR=1.7 (95% CI: 1.2-2.4) times greater than the risk of SCC amongst those at low risk, in logistic regression following adjustment for just gender (as age group was considered when allocating study participants to high or low risk groups).

Additional logistic models found SCC risk to be slightly higher among those who were JC seropositive compared to seronegative when restricted to participants ≥65 years of age (OR=1.6, 95% CI: 1.1-2.5) or MCV seronegative (OR=1.9, 95% CI: 1.0-3.6). Logistic models created to mimic the best MDR model consisted of 4 analyses that stratified participants by a combination of median age (</≥65 years) and MCV serostatus (seropositive or seronegative) (Supplemental Figure 3.5). However, in a logistic regression model restricted to just those participants who were MCV seropositive, no significant interaction between JC serostatus and median age was found (*P*-for-interaction=0.34). The most relevant comparison revealed a pattern in risk supported by the top MDR result, whereby restriction to those who were MCV seropositive and <65 years (OR=0.8, 95% CI: 0.4-1.6), and MCV seropositive and ≥65 years (OR=1.5, 95% CI: 0.9-2.5), showed a difference in the odds of SCC with JC seropositivity.

### 3.4.4 Immuno-genotype findings

Of the 713 case and control participants, 615 (86.2%) had *HLA-C* genotyping data, 652 (91.4%) had *HLA-C -35 kb* genotyping data, 654 (91.7%) had *NKG2Da* genotyping data, and 638 (89.5%) had *NKG2Db* genotyping data. No appreciable differences were noted in the characteristics of the individuals for whom we did not obtain immuno-genotyping data. The allele frequencies were in Hardy-Weinburg equilibrium for all study participants, cases separately, and controls separately.

No association between *HLA-C*, the *HLA-C -35 kb* expression SNP, *NK2Da*, or *NKG2Db* genotype and SCC case or control status was observed, nor was an association found following adjustment for age group and gender. Amongst all study participants, there were some differences in the frequency of *HLA-C* genotypes by skin color; *HLA-C -35 kb* genotypes by number of lifetime painful sunburns; and *NKG2Db* genotypes and smoking status and glucocorticoid use (Supplemental Table 3.4). Amongst the controls, WU serostatus was found to differ by *NKG2Da* and *NKG2Db* genotype; MCV serostatus to differ by *HLA-C* genotype; and HPyV6 serostatus to differ by *HLA-C -35 kb* genotype (Supplemental Table 3.5).

Overall, there was little to no evidence of effect modification by *HLA-C* or *NKG2D* genotype on SCC risk by PyV VP1 serostatus (Supplemental Figures 3.6-3.9) or quartiles of seroreactivity (Supplemental Tables 3.6-3.9) for our limited subset of participants for whom we had genotype data. While generally null, when stratified by *HLA-C* genotype and amongst those with the C2 variant, the risk of SCC with HPyV9 seropositivity was 3.9 (95% CI: 1.1-13.7, *P*=0.03) times greater than the risk of SCC amongst those who were HPyV9 seronegative, following adjustment for age group and gender (Supplemental Figure 3.6).

## 3.5 DISCUSSION

We measured antibodies against the first ten discovered human PyVs in a large, population-based case-control study to investigate the relation between multiple PyVs and SCC incidence. No clear associations were observed for most PyV types and SCC risk. However, we observed an intriguing elevated odds ratio for SCC associated with JC seropositivity that approached statistical significance, and a trend in risk by quartiles of JC seroreactivity. While no clear associations were detected, we had limited statistical power for many of these analyses due to the high PyV seroprevalence.

Several epidemiologic studies have raised the possibility of an oncogenic role for PyV infection in SCC development. Among 173 SCC cases and 300 controls in a clinic-based case-control study conducted in Florida, USA, MCV seropositivity was related to an increased risk for SCC (OR=1.58, 95% CI: 0.96-2.60), and the association was strongest among those with tumors containing MCV DNA (OR=2.49, 95% CI: 1.03-6.04).[54] We did not observe evidence of an association with MCV seropositivity. An association of a similar magnitude with JC seropositivity was also observed in the Florida study (OR=1.40, 95% CI: 0.89–2.20). However, unlike our study, there was no trend by quartiles of JC seroreactivity (*P*-for-trend=0.44).[54]

In our case-control study design, we are unable to address the issue of 'reverse causality' common to studies aiming to elucidate the etiology of KC at or following diagnosis. Namely, it is possible antibodies measured at the time of SCC diagnosis do not accurately reflect antibodies that would have been circulating in the early stages of carcinogenesis. Indeed, a large, prospective, Swedish study which drew upon samples donated to a biobank and cancer registry data found no association between SCC (OR=1.0, 95% CI: 0.8-1.4) or BCC (OR=0.9, 95% CI: 0.8-1.1) and prediagnostic JC seropositivity, with blood samples collected at least 1 month prior to diagnosis.[339] However, the primary hypothesis of the Swedish study was an association between HPV and SCC or BCC, and thus used a multiplex assay optimized for the detection of antibodies against various HPV types, with JC included as a specificity control.[339] A discrepancy in results may be attributed to both differing populations and assay optimization for human PyV types. Thus, further prospective data that include repeated measurements over time from each participant are needed.

Increased incidence of cancers in immunosuppressed patients – such as skin cancers,[84] suggests a possible viral etiology that may be attributed to impaired immune surveillance of

infections with oncogenic potential. Moreover, immune suppression could lead to viral reactivation and higher levels of antibodies, thus confounding any observed relationship between PyVs and SCC. A greater than expected proportion of transplant recipients are seropositive for KI and WU,[350, 351] and MCV sequences are also more frequently found in the skin tumors of immunocompromised patients than those from immunocompetent individuals.[85] TSV seroprevalence is likewise higher in renal transplant patients when compared to healthy participants,[50] and has been causally linked to a rare disease of abnormal maturation of the hair follicles characterized by spiny lesions on the skin.[10, 11] HPyV9 was first isolated from a kidney transplant recipient[28] and has been found at higher seroprevalence in immunocompromised patients.[305] HPyV10 was isolated from a patient with warts, hypogammaglobulinemia, infections, and myelokathexis (WHIM) syndrome,[30] an inherited immune deficiency with increased susceptibility to HPV-induced warts and cancers.[352, 353] BK and JC are known to reactivate under conditions of immunosuppression.[7, 354] Indeed, JC seropositivity has been used as an indicator of immune status in prior studies.[54] In a small, nested case-control study among transplant recipients, no excess risks for SCC following organ transplantation were found with antibody levels against PyVs (specifically, JC, KI, WU, MCV, HPyV6, and HPyV7) in serum drawn immediately prior to surgery (n=149 SCC cases and 290 controls).[87] We observed increased odds of SCC associated with increasing JC antibody levels, which may reflect the detection of JC reactivation arising from immune dysregulation associated with skin cancer. Therefore, our findings of an association between JC serostatus and SCC could represent the predisposing role of immunodeficiency, rather than the virus itself. With limited statistical power, there was weak evidence that the association we found between SCC and JC seropositivity may be slightly stronger among those with prolonged oral glucocorticoid use, raising the possibility of

effect modification rather than confounding by immunosuppression. However, further studies are needed.

***NK cells are often the immune system's first line of defense against transformed or infected host cells.[355] NK cells operate through their KIR receptors which are licensed to bind specific diverse and polymorphic human leukocyte antigen (HLA) class I ligands on target cells, resulting in cytolysis if the correct "signals" are received by the NK cell.[355] The diverse and polymorphic nature of the KIR gene family is thought to drive NK cell killing phenotypes and inter-individual variation in these interactions (reviewed in Middleton et al., 2010).[356] Inter-individual variation in the KIR and HLA loci have been associated with susceptibility to numerous infectious and autoimmune diseases, such as HIV[357] and hepatitis C,[358] as well as cancers (e.g., NMSC).[343, 344, 345, 359] However, we found no evidence of effect modification of the relation between PyV serostatus and SCC risk by HLA-C or NKG2D genotype in our small study. The role of underlying immuno-genotype in SCC risk requires additional study in larger populations as it was interesting that the proportion of CD8+ T cells varied with HPyV9 serostatus in our previously published work amongst the controls,[341] and SCC risk among those who were HPyV9 seropositive compared to seronegative was higher for those who had the C2 genotype for HLA-C, as HLA-C presents antigens to CD8+ T cells to generate a response.[360, 361]***

MCV, TSV, HPyV6, and HPyV7 show evidence of skin tropism,[3] yet we found no association with SCC risk. We could not assay for the presence of PyV DNA within SCC tumor tissues, which is a more direct indication of a potential association between PyVs and SCC. In the Florida study, MCV TAg DNA was isolated from ~40% of SCC tumors from patients, but they did not find BK, JC, KI, or WU TAg DNA sequences within SCC tumors.[54] Other small studies have failed to detect DNA from JC in skin swabs[362] or skin biopsies from normal skin,[363]

common warts, Bowen's disease,[364] BCC,[365] SCC,[363, 365] melanoma,[82, 363] or cutaneous B cell lymphomas.[366] However, JC TAg sequences were detected in Kaposi's sarcoma skin lesions, and has been amplified from 16% of healthy skin tissues,[81] suggesting that these viruses can inhabit the skin even without being skin tropic[3] or the skin acting as a suitable host for viral propagation.[367] A potential etiologic role of JC in the absence of JC DNA within tumors remains unexplained and warrants additional study.

All human PyVs discovered thus far encode proteins that allow them to act as oncoviruses. Indeed, the virus' name is derived from the Greek *poly* meaning 'many' and *oma* referring to their induction of tumors in murine models.[1] Truncating mutations in the large TAg of integrated MCV are characteristic of MCC tumors, with domains required for Rb-induced cell transformation preserved, while those for viral replication and p53 binding[12] are eliminated resulting in no interaction with p53,[368] thereby enhancing the likelihood of cell survival. Truncating mutations of the TAg have not been described in other human PyVs. Consequent to its transforming potential, the large TAg has been considered the main oncoprotein in PyVs.[3] The small TAg and agnoprotein expressed by some viruses also exhibit oncogenic properties.[3] The late regions of BK and JC encode an agnoprotein which exerts a tumorigenic effect through cell cycle dysregulation, interference with DNA repair processes, and increased chromosome instability.[67, 68] Similar to SV40,[369] BK, JC, and MCV encode a miRNA that downregulates large TAg expression, which may allow the virus to escape the immune system.[70] Thus, there is some evidence for a putative mechanistic role for PyVs in human carcinogenesis, although further research is needed.

***It has been theorized that PyVs may serve as cofactors for oncoviruses by acting on common tumorigenic targets, perturbing the cell cycle, assisting with immune evasion, or trans-activating the***

*promoters of co-infecting viruses – although evidence of such an interaction is lacking.[333] Interactions between PyVs and other viruses have been documented, such as the in vitro cooperation between JC and the HIV-1 regulatory protein, Tat, that enhances JC transcription in glial cells.[370, 371] Additionally, interactive effects between a murine polyomavirus strain with Moloney murine leukemia virus have been shown to result in stunted growth only in co-infected animals, possibly through a proinflammatory cytokine pathway.[372] In our study, we were only able to assess concomitant seropositivity to β HPVs, and did not detect any consistent interactions with PyV types and SCC risk. MDR analysis identified a strong synergistic interaction between JC and age group, and JC and age group with MCV, as being related to SCC. Evidence of synergy between JC and MCV implies this interaction provides information on SCC case or control status that cannot be obtained from individual inspection of the association between SCC and PyV seropositivity, i.e., in logistic regression models. MCV has been described as a "cutaneous" PyV,[3] and it was interesting that it appeared to be related to a NMSC outcome in MDR as it has been previously implicated as a etiologic agent in MCC. The inclusion of JC – a virus that reactivates under conditions of immunosuppression and has been used as an indicator of immune status,[54] could represent the predisposing role of immunodeficiency in the association between SCC and cutaneous PyVs. Although the MDR results had poor model statistics, concordance with a series of logistic regression models that aimed to mimic the top MDR model leant support to the MDR findings.*

### 3.5.1 Strengths and limitations of the study

A strength of our study was the large number of histologically confirmed cases of incident invasive SCC identified through active population-based surveillance, along with controls derived from the general USA population. This type of study design decreases the opportunity for selection bias, and is more generalizable than clinic- or hospital-based case-control studies. Still,

the possibility of selection bias and residual confounding cannot be excluded, and the generalizability to non-white populations is limited due to the study's location in an almost exclusively white USA population. The use of multiplex serology to comprehensively measure a wide range of human PyVs was an additional study strength. The GST capture of recombinantly expressed VP1 capsid proteins has been found to be a reliable technique to assess PyV seroreactivity and has been used as a marker of PyV infection in prior studies.[39, 53, 54] Our blood samples were collected following the skin cancer diagnosis, and thus PyV infection may have occurred during or following tumor development, and the influence of disease state on susceptibility to PyV infection and immune response cannot be disentangled from these data. We are unable to assess temporality or reverse causality, so the direction of any observed associations cannot be determined. PyV seropositivity does not necessarily correspond to the presence of PyV DNA in tumor tissues and seroreactivity is an indirect measure of infection. However, MCV viral load and antibody titer have been shown to have a strong positive monotonic correlation,[49] with higher MFI values corresponding to MCV DNA-positive SCC tumor tissues.[54] Lastly, given the high seroprevalence of PyVs in the study population, power to detect differences in SCC risk for some PyVs was limited.

### 3.5.2 Conclusions

In this population-based case-control study, we examined the association between the first 10 discovered human PyVs and incidence of SCC using multiplex serology. Our findings, though limited, provide some support for the possibility that specific PyV types may play a role in the occurrence of SCC in the United States population, but not PyV seropositivity in general.

# Chapter 4

# Prospective study of human polyomaviruses and risk of cutaneous squamous cell carcinoma in the United States

*Published as: Anala Gossai, Tim Waterboer, Heather H. Nelson, Jennifer A. Doherty, Angelika Michel, Martina Willhauck-Fleckenstein, Shohreh F. Farzan, Brock C. Christensen, Anne G. Hoen, Ann E. Perry, Michael Pawlita, and Margaret R. Karagas. "Prospective study of human polyomaviruses and risk of cutaneous squamous cell carcinoma in the United States." Cancer Epidemiology, Biomarkers & Prevention (2016): cebp-1111.*

## 4.1 ABSTRACT

Merkel cell polyomavirus (PyV) is causally related to Merkel cell carcinoma, a rare skin malignancy. Little is known about the serostability of other PyVs over time, or associations with cutaneous squamous cell carcinoma (SCC). As part of a US nested case-control study, antibody response against the PyV VP1 capsid proteins of BK and JC was measured using multiplex serology on 113 SCC cases and 229 gender, age, and study center-matched controls who had a prior keratinocyte cancer. Repeated serum samples from controls, and both pre- and post-diagnosis samples from a subset of SCC cases were also tested. Odds ratios (OR) for SCC associated with seropositivity to each PyV type were estimated using conditional logistic regression. Among controls, BK and JC seroreactivity was stable over time, with intraclass correlation coefficients of 0.86 for BK and 0.94 for JC. Among cases, there was little evidence of seroconversion following SCC diagnosis. JC seropositivity prior to diagnosis was associated with an elevated risk of SCC (OR=2.54, 95% CI: 1.23-5.25), and SCC risk increased with increasing quartiles of JC (*P*-for-trend=0.004) and BK (*P*-for-trend=0.02) seroreactivity. PyV antibody levels were stable over time and following an SCC diagnosis. A history of PyV infection may be involved in the occurrence of SCC in a population at high risk for this malignancy. A single measure of PyV seroreactivity appears a reliable indicator of long-term antibody status, and PyV exposure may be a risk factor for subsequent SCC.

## 4.2 INTRODUCTION

The human polyomavirus (PyV) is a non-enveloped virus with an icosahedral capsid containing a circular double-stranded DNA genome.[2, 3] The genome of the *Polyomaviridae* family encodes three capsid proteins (VP1, VP2, VP3), as well as small and large T antigens (TAg).[1, 2] The large TAg (LTAg) has oncogenic potential,[2, 15, 16] but there is limited support for carcinogenesis in humans.

PyV infection rates vary between populations and viral types, and seroprevalence usually increases with age.[37, 41, 42, 46, 48, 305] Little is known about intraindividual PyV antibody stability over time in the general population,[373] but repeated measures of PyV seroreactivity collected from individuals with a compromised immune system[34, 316, 374, 375, 376, 377, 378, 379] suggest antibody levels may be consistent longitudinally.

PyV infections are ubiquitous within human populations.[4] Among immunosuppressed patients, BK virus is the etiologic agent of polyomavirus-associated nephropathy and cystitis,[7, 8] and John Cunningham virus (JC) reactivation has been linked to progressive multifocal leukoencephalopathy.[7, 8] Merkel cell polyomavirus (MCV) containing mutations in LTAg[12] has been established as a causal factor for Merkel cell carcinoma (MCC) – a rare but aggressive skin cancer.[12, 13, 14, 24]

Cutaneous squamous cell carcinoma (SCC) arises from epithelial keratinocytes,[101] and is a common malignancy with increasing incidence rates reported in the United States.[102, 103, 120, 380, 381] SCC etiology is largely attributed to ultraviolet radiation,[96, 170] but other risk factors including immunosuppression[84, 213, 214] raise the possibility of a viral etiology. An oncogenic role for PyV infection in the development of SCC has been hypothesized in recent epidemiologic studies.[54] A clinic-based case-control study from Florida, USA, found an increased SCC risk associated with

antibodies against MCV assessed at the time of diagnosis.[54] Conversely, a case-control study conducted among organ transplant recipients found no association between SCC development following transplant surgery and seropositivity to multiple PyV types prior to transplantation.[87] There are limited prospective studies assessing whether past virus exposure predicts risk of future SCC development.

Therefore, using data and stored serum samples from patients with a prior history of keratinocyte cancer (KC) enlisted in a USA skin cancer prevention trial, we performed a longitudinal analysis of the presence and stability of antibodies against human PyV types BK and JC, and conducted a nested case-control study to investigate the role of polyomaviruses on subsequent SCC incidence.

## 4.3 METHODS

### 4.3.1 Patient population and parent study design

We derived our study group from the *Skin Cancer Prevention Study* – a multicenter, randomized clinical trial conducted in the United States from 1980 to 1989 to test the efficacy of oral β-carotene supplementation in the prevention of KC among persons with a prior history of this malignancy[159, 382] (Supplemental Figure 4.1). The trial methods and study participants have been described in detail elsewhere.[159, 382, 383] Briefly, patients were 35-<85 years of age and had had at least one biopsy-proven SCC or basal cell carcinoma (BCC) removed since January 1, 1980. Of the 5,232 potentially eligible patients identified through a review of dermatopathology reports in the four clinical centers, 1,805 (34.5%) fulfilled the trial criteria, and were subsequently enrolled for randomization to receive either β-carotene or a placebo.

Upon entry to the trial, patients completed a questionnaire regarding individual characteristics, including age, hair and eye color, cigarette smoking, sun exposure, and medical

history (e.g., vitamin use). A dermatological evaluation determined each patients' skin type (i.e., tendency to sunburn, extent of solar damage), and the histological type and number of previous KC diagnoses was documented from patients' medical records.

Follow-up consisted of an interval questionnaire mailed every 4 months. A dermatological examination was conducted at enrollment and annually thereafter, during which a 20 mL blood specimen was collected and stored in heparinized vacuum tubes at -75°C until analysis. The appearance of new, primary skin cancers was monitored, and microscopic slides of suspected cancerous lesions were re-reviewed by a dermatopathologist at the study coordinating center for independent validation. The primary trial end point was the first occurrence of a new basal or squamous cell carcinoma after randomization. Follow-up for each patient continued for 5 years or until September 30, 1989 when the treatment phase of the trial ended. All participants provided informed consent in accordance with the Committee for the Protection of Human Subjects at Dartmouth.

### 4.3.2 Nested case-control study design

We conducted a nested case-control study within this intervention trial to examine the risk of subsequent incident SCC (i.e., the first new occurrence of a nonrecurrent squamous cell skin cancer following randomization) in relation to PyV infection status prior to diagnosis of this new SCC (hereafter referred to as the 'prediagnostic' or 'prior to diagnosis' time period) among patients with a history of KC. Of the 1,805 patients enrolled in the trial, 132 (7.3%) developed a new, nonrecurrent SCC during the follow-up period (hereafter referred to as a 'case'). For each of these case patients, 2 controls were randomly selected from among patients who, up until or during the study year the SCC case was diagnosed, *(a)* had been actively followed, and *(b)* had not developed an incident and nonrecurrent SCC (hereafter referred to as a 'control'). Controls

were pair-matched to cases on gender, age (<45, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, and 75-84 years), and study center (Hanover, NH; Minneapolis, MN; Los Angeles, CA; and San Francisco, CA). Controls were assigned a reference date corresponding to the case's diagnosis date to whom they were matched.

We aimed to analyze the baseline (pre-randomization) serum sample for the determination of PyV seroreactivity for each case and control included in our study sample. If the baseline sample was unavailable, we tested the earliest blood sample collected – provided it was drawn prior to the diagnosis date for SCC cases or the reference date for controls.

### 4.3.3 Repeated measures

A subset of 89 cases had both a pre-diagnosis blood, and a post-diagnosis blood drawn nearest to, but following, the diagnosis date of the new SCC occurrence. We further investigated the stability of PyV antibodies over time in serial serum samples through a longitudinal serological study conducted among controls included in the nested case-control study. A total of 895 serum samples from 229 controls were available for serologic analysis. Control participants were included in this longitudinal sub-study if they had ≥2 serum samples collected during the follow up period of 6 years.

### 4.3.4 Human polyomavirus serology

Serum samples masked to case-control status were shipped on dry ice to the German Cancer Research Center (DKFZ; Heidelberg, Germany) for analysis. Serum samples were assayed for antibodies against the immunodominant VP1 capsid protein[53] of two human-associated PyV types: BK[21] and JC.[20] PyV seroreactivity was determined using a multiplex antibody detection approach based on a glutathione S-transferase (GST) capture enzyme-linked immunosorbent assay method in combination with fluorescent bead technology (Luminex Corp., Austin, Texas,

USA).[55, 58] Antigen preparation and techniques used for PyVs[39, 53, 54] closely follow methods applied to human papillomaviruses (HPV) as described previously.[55, 56] Although the multiplex technology assayed for other viruses simultaneously (i.e., HPVs), BK and JC were the only PyVs included in the assay.

Seroreactivity against the PyV VP1 antigens was expressed as the median fluorescence intensity (MFI) of 100+ beads of the same internal color.[58] MFI values reflect viral load,[49] as well as antibody affinity, titer, and reactivity as determined by dilution series.[60] Standard cut points to define seropositivity were chosen for each PyV by visual inspection of frequency distribution curves (percentile plots) for the inflection points of all sera tested, as done in prior studies.[14, 39, 56, 62] The standard cutoff value for VP1 seropositivity was 400 MFI units. To evaluate the robustness of PyV VP1 seroprevalence and odds ratio (OR) estimates for SCC by PyV seropositivity, we used a sliding cut point between 250 and 550 MFI units. We ultimately used the standard cut points in all analyses as PyV seroprevalence (Supplemental Figure 4.2) and OR estimates (from conditional and unconditional logistic regression models; Supplemental Figures 4.3-4.4) were insensitive to cut point definition.

### 4.3.5 Statistical analyses

Statistical analyses were performed in R version 3.1.0. All statistical tests were two-sided, and statistical significance assessed at the $\alpha=0.05$ level. Individual characteristics of SCC cases and controls were compared using the $X^2$ test (for categorical variables, i.e., gender, randomization arm, study center, previous skin cancers, cigarette use, skin sun sensitivity, occupational sun exposure, eye color, hair color, vitamin use) or Fisher's exact test (for categorical variables with small strata containing $\leq$10 persons, i.e., BMI category, extent of UV skin damage, sun bathing), and Wilcoxon rank sum test (for continuous variables, i.e., age). Among controls, the

seroprevalence of each PyV type was examined for both PyV seropositivity overall, and by age groups, using binary MFI cut points. Additionally, we tested the association between various individual characteristics in relation to PyV seropositivity within controls. We used the continuous MFI values from controls to compute Spearman rank correlation coefficients (ρ) between both PyVs assayed.

Within controls, we investigated intraindividual changes in PyV seroreactivity over time using repeated measures taken after baseline by calculating the intraclass correlation coefficient (ICC)[384] for continuous MFI values, and also stratified analyses by randomization arm of the original trial (i.e., treatment or placebo). Additionally, to determine serostatus (positive versus negative) stability over time, we defined control participants as being seropositive at all time points (stably seropositive), seronegative at all time points (stably seronegative), seroconverting (change from seronegative to seropositive over time), seroreverting (change from seropositive to seronegative), or fluctuating between seropositive and seronegative (fluctuating), as has been done previously.[316, 373, 385] The likelihood of seroconversion following SCC diagnosis was evaluated with the kappa (κ) statistic among cases.[386, 387, 388]

We used both conditional (as there was one-to-two pair-matching between cases and controls in the original trial) and unconditional logistic regression (while adjusting for the matching factors: age, gender, and study center) to calculate the ORs and 95% confidence intervals (CI) for the development of a new SCC by VP1 seropositivity compared to seronegativity for each PyV type in the baseline or earliest blood sample collected. Quartiles of seroreactivity based on the control distributions of continuous MFI values were created for each PyV, and associated with SCC by comparing the second, third, and fourth quartiles to the first (lowest) quartile in conditional and unconditional logistic regression models. Tests-for-trend

were conducted by including an ordinal variable in the logistic model in place of the categorical quartile variable. Using the study sample for unconditional models, generalized additive logistic models (GAM) were fit to evaluate deviations from linearity in risk of SCC by the continuous MFI values from the earliest blood samples.[389, 390] The smoothed (nonparametric) component of the binomial GAMs was PyV seroreactivity, and adjustment was made for the unsmoothed matching factors. Models were not further adjusted for potentially confounding covariates, as no sociodemographic or individual characteristic was found to both be associated with PyV serostatus and a risk factor for SCC development in our study group. We assessed the potential modifying effects of the assigned randomization arm from the original trial, having had a prior SCC, and having had a prior BCC in stratified analyses. We also performed stratified analyses by smoking status, UV skin damage, skin sun sensitivity, and hair color.

## 4.4 RESULTS

### 4.4.1 Patient characteristics

We tested a prediagnostic serum sample for 113 (85.6%) of the 132 SCC cases, and 229 (86.7%) of the 264 controls, for the nested case-control study. Pair-matched sets with at least one measured control directly matched to a case consisted of 111 (84.1%) SCC cases and 195 (73.9%) controls. Tested serum samples for all 342 study participants were drawn by the first year after baseline on 93.3% of both cases and controls, and samples were taken 15 days to 5.3 years before the diagnosis or reference date (median=2.2 years, interquartile range (IQR)=1.3-3.3 years). Among 210 controls (excluding 19 controls who only had a single sample collected), we performed repeated serologic analysis on 876 (97.9%) of the 895 serum samples drawn during the follow-up period for the longitudinal study. Controls had 2 to 8 serial samples with a median of 4 samples per participant, taken 11 days to 4.2 years apart (median number of years

between repeated measures=1.0 years, IQR=1.0-1.1 years). We analyzed 85 (95.5%) of the subset of 89 SCC cases with both pre- and post-diagnosis serum samples, and the post-diagnostic samples were obtained 17 days to 1.7 years following the diagnosis date (median=0.7 years, IQR=0.6-0.7 years).

Participants in this study ranged in age from 35 to 84 years (median age of 67 years) upon study entry, and 88.1% were men. Cases and controls were balanced with respect to age, gender, and study center through the matched design. Compared to controls, SCC cases were more likely to have had ≥2 previous KCs, be current or former cigarette smokers, have skin that always or usually burned with sun exposure, be blonde or red-haired, and have moderate to severe actinic damage (Table 4.1).

Table 4.1. Distribution of selected baseline characteristics among cutaneous squamous cell carcinoma (SCC) cases and controls from the Skin Cancer Prevention Study (n=342).[a]

| Characteristic | SCC Cases (n=113), No. (%) | Controls (n=229), No. (%) |
|---|---|---|
| Gender | | |
| Male | 100 (88.5) | 201 (87.8) |
| Female | 13 (11.5) | 28 (12.2) |
| Median age, SD (years) | 68 (8.0) | 67 (8.2) |
| Randomization arm in RCT | | |
| Treatment | 66 (58.4) | 115 (50.2) |
| Placebo | 47 (41.6) | 114 (49.8) |
| Study center[b] | | |
| DHMC | 24 (21.2) | 43 (18.8) |
| UCLA | 32 (28.3) | 65 (28.4) |
| UCSF | 25 (22.1) | 56 (24.4) |
| UMN | 32 (28.3) | 65 (28.4) |
| Previous skin cancers | | |
| 1 | 24 (21.2) | 102 (44.5)*** |
| 2 | 21 (18.6) | 41 (17.9) |
| 3 | 12 (10.6) | 20 (8.7) |
| 4-5 | 26 (23.0) | 36 (15.7) |
| 6-9 | 12 (10.6) | 17 (7.4) |
| ⩾10 | 18 (15.9) | 11 (4.8) |
| Cigarette use | | |
| Never smoked | 28 (24.8) | 91 (39.7)** |
| Former smoker | 62 (54.9) | 111 (48.5) |
| Current smoker | 23 (20.3) | 27 (11.8) |
| Body mass index (kg/m$^2$) | | |
| Underweight <18.5 | 1 (0.9) | 2 (0.8) |
| Normal 18.5-24.9 | 50 (44.2) | 98 (42.8) |
| Overweight 45.0-29.9 | 47 (41.6) | 108 (47.2) |
| Obese >30.0 | 12 (10.6) | 15 (6.6) |
| Skin sun sensitivity | | |
| Always or usually burns | 72 (63.7) | 110 (48.0)** |
| Burns moderately or minimally | 41 (36.3) | 118 (51.5) |
| Extent of UV skin damage | | |
| Mild | 9 (8.0) | 62 (27.1)*** |
| Moderate | 71 (62.8) | 134 (58.5) |
| Severe | 32 (28.3) | 31 (13.5) |
| Sun bathed (hours) | | |
| Never | 40 (35.4) | 62 (27.1) |
| 0-200 | 18 (15.9) | 63 (27.5) |
| 200-400 | 25 (22.1) | 54 (23.6) |
| 400-600 | 21 (18.6) | 33 (14.4) |
| >600 | 7 (6.2) | 17 (7.4) |
| Occupational sun exposure (years) | | |
| 0-7 | 41 (36.3) | 78 (34.1) |
| 7-20 | 31 (27.4) | 64 (27.9) |
| 21-40 | 28 (24.8) | 40 (17.5) |
| >40 | 13 (11.5) | 46 (20.1) |
| Eye color | | |
| Blue, green, gray, hazel | 97 (85.8) | 185 (80.8) |
| Brown, black | 16 (14.2) | 44 (19.2) |
| Hair color | | |
| Blonde, red | 49 (43.4) | 61 (26.6)** |
| Brown, black | 64 (56.6) | 64 (27.9) |
| Vitamin use | | |
| No | 67 (59.3) | 128 (55.9) |
| Occasional | 15 (13.3) | 37 (16.2) |
| Daily | 24 (21.2) | 60 (26.2) |

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. $P$ values obtained from $X^2$, Fisher's exact, or Wilcoxon rank sum test (as appropriate) comparing sociodemographic and skin cancer risk factors between SCC cases and controls.

[a] Numbers may not sum to the overall total due to missing data.

[b] This multicenter study was conducted at sites in California (University of California at Los Angeles School of Medicine (UCLA); University of California Medical School, San Francisco (UCSF)), Minnesota (University of Minnesota Schools of Medicine and Public Health, Minneapolis (UMN)), and New Hampshire (Dartmouth-Hitchcock Medical Center, Hanover (DHMC)), USA.

### 4.4.2 BK and JC antibody status over time among controls

Among the baseline or first measured samples in controls, the overall seroprevalence was 96.9% for BK and 77.3% for JC (Figure 4.1). Seroprevalence was constant across age groups for BK, and increased with age group for JC (*P*-for-trend=0.02). Sociodemographic and individual characteristics at study entry were not related to BK or JC serostatus (Supplemental Table 4.1). We did not find correlations or evidence of cross-reactivity between the VP1 capsid proteins of BK and JC for the earliest prediagnostic samples collected (ρ=0.06, *P*=0.09) or for the repeated measures (ρ=0.06, *P*=0.07).

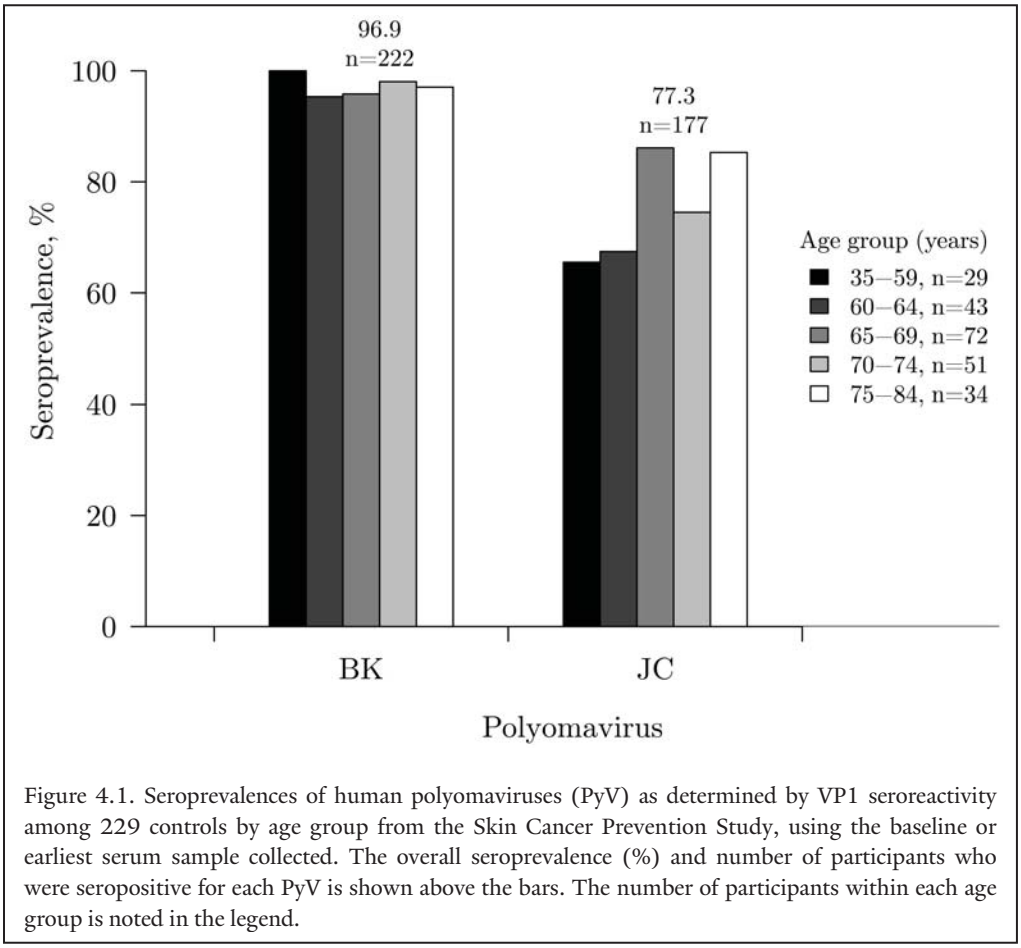



Figure 4.1. Seroprevalences of human polyomaviruses (PyV) as determined by VP1 seroreactivity among 229 controls by age group from the Skin Cancer Prevention Study, using the baseline or earliest serum sample collected. The overall seroprevalence (%) and number of participants who were seropositive for each PyV is shown above the bars. The number of participants within each age group is noted in the legend.

BK and JC seroreactivity remained stable over time within controls in the repeated serum measures (Table 4.2). The intraclass correlation coefficient was 0.86 for BK and 0.94 for JC, and no difference was found in ICC estimates when stratified by randomization arm of the original trial. For BK, 95.7% were stably seropositive, and 2.9% remained seronegative. There were no seroreversions, and <1% seroconverted (0.9%) or had fluctuating antibody levels (0.5%). For JC, 74.8% were stably seropositive, and 20.9% remained seronegative; 2.9% seroconverted, 0.9% seroreverted, and 0.5% had fluctuating antibody levels.

Table 4.2. Serostability for BK and JC human polyomavirus (PyV) seropositivity [a] in samples collected longitudinally over time among 210 [b] controls from the Skin Cancer Prevention Study.

| PyV | Controls Serostability (n=210), No. (%) | | | | |
|---|---|---|---|---|---|
| | Stably seropositive | Stably seronegative | Seroconversion | Seroreversion | Fluctuating |
| BK | 201 (95.7) | 6 (2.9) | 2 (0.9) | None | 1 (0.5) |
| JC | 157 (74.8) | 44 (20.9) | 6 (2.9) | 2 (0.9) | 1 (0.5) |

[a] PyV infection was determined using seropositivity for the VP1 protein.
[b] Not including 19 controls who only had a single sample collected. Only controls with ≥2 repeated serum samples were included.

### 4.4.3 Pre- versus post-diagnostic BK and JC serostatus

We compared PyV serostatus prior to and closely following the first occurrence of an incident SCC diagnosis among a subset of SCC cases (Table 4.3). One (1.2%) case who was BK seronegative prior to diagnosis remained seronegative following SCC diagnosis, and 84 (98.8%) cases who were seropositive prior to diagnosis remained so following diagnosis. For JC, we found 10 (11.8%) cases to remain seronegative, and 74 (87.1%) cases to remain seropositive, prior to and following diagnosis; only 1 (1.2%) case who was seronegative seroconverted to JC seropositive following SCC diagnosis. As a measure of agreement, the kappa statistic (κ) was close to 1 for both BK (κ=1.0, 95% CI: 0.79-1.0) and JC (κ=0.95, 95% CI: 0.73-1.0).

Table 4.3. BK and JC human polyomavirus (PyV) serostatus [a] prior to cutaneous squamous cell carcinoma diagnosis (SCC), and post SCC diagnosis, among 85 cases from the Skin Cancer Prevention Study.

| PyV serostatus pre-diagnosis | PyV serostatus post-diagnosis (n=85) | |
|---|---|---|
| | Seronegative, No. (%) | Seropositive, No. (%) |
| BK | | |
| Never seropositive | 1 (1.2) | 0 (0) |
| Ever seropositive | 0 (0) | 84 (98.8) |
| JC | | |
| Never seropositive | 10 (11.8) | 1 (1.2) |
| Ever seropositive | 0 (0) | 74 (87.1) |

[a] PyV infection was determined using seropositivity for the VP1 protein.

We assessed risk of having a new SCC in relation to prediagnostic PyV seroreactivity among patients with a history of KC (Table 4.4). In the conditional logistic regression models of the matched cases and controls, an increased risk of SCC was specifically observed with seropositivity for JC (OR= 2.54, 95% CI: 1.23-5.25), and a positive trend in SCC risk was associated with increasing MFI quartiles of JC seroreactivity (*P*-for-trend=0.004). The highest versus lowest quartile of JC seroreactivity was associated with a 3-fold greater risk of SCC (OR=3.13, 95% CI: 1.51-6.49). In conditional analyses, a positive association with SCC risk was also observed with seropositivity for BK (OR= 3.90, 95% CI: 0.48-31.96) with limited statistical power, and there was evidence of an increasing trend in SCC risk with increasing quartiles of seroreactivity (*P*-for-trend=0.02). Similar estimates were obtained using unconditional logistic regression models that adjusted for the matching factors: age, gender, and study center (Supplemental Table 4.2). The binomial GAMs suggested an increasing probability of SCC with increasing continuous MFI values for both BK (*P*=0.02) and JC (*P*=0.003) (Figure 4.2), with some deviations from linearity and statistical imprecision (especially for BK). Analyses stratified by randomization arm, prior SCC status (220 controls and 107 cases with no prior SCC presented with their first incident SCC), prior BCC status (Supplemental Figure 4.5),

smoking status, UV skin damage, and hair color (Supplemental Figure 4.6) resulted in similar OR estimates as the main analyses but were limited by the markedly reduced sample sizes. The OR associated with JC was higher among those with a skin type that tended to burn, but this was based on small stratum sizes and had wide CIs (Supplemental Figure 4.6).

Table 4.4. Conditional odds ratios [a] (95% confidence intervals) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) [b] type and quartiles [c] of PyV seroreactivity at baseline among cases and matched controls (n=306) from the Skin Cancer Prevention Study.

| | | SCC Cases (n=111) | |
|---|---|---|---|
| **PyV seroreactivity (MFI units)** | **Controls (n=195), No. (%)** | **No. (%)** | **OR (95% CI)** |
| BK | | | |
| Seronegative | 7 (3.6) | 1 (0.9) | 1.00 (referent) |
| Seropositive | 188 (96.4) | 110 (99.1) | 3.90 (0.48-31.96) |
| Quartile 1 | 54 (27.7) | 26 (23.4) | 1.00 (referent) |
| Quartile 2 | 48 (24.6) | 14 (12.6) | 0.65 (0.31-1.37) |
| Quartile 3 | 46 (23.6) | 32 (28.9) | 1.67 (0.86-3.23) |
| Quartile 4 | 47 (24.1) | 39 (35.1) | 1.89 (0.96-3.73) |
| *P* for trend [d] | | | 0.016 |
| JC | | | |
| Seronegative | 45 (23.1) | 13 (11.7) | 1.00 (referent) |
| Seropositive | 150 (76.9) | 98 (88.3) | 2.54 (1.23-5.25) |
| Quartile 1 | 50 (25.6) | 18 (16.2) | 1.00 (referent) |
| Quartile 2 | 49 (25.1) | 26 (23.4) | 1.60 (0.76-3.39) |
| Quartile 3 | 51 (26.1) | 21 (18.9) | 1.10 (0.50-2.44) |
| Quartile 4 | 45 (23.1) | 46 (41.4) | 3.13 (1.51-6.49) |
| *P* for trend [d] | | | 0.0039 |

[a] OR=odds ratios obtained from conditional logistic regression analysis, CI=confidence interval.
[b] PyV infection was determined in the baseline or earliest serum sample collected using seropositivity for the VP1 protein.
[c] Controls may not be evenly distributed within quartiles due to uneven data distribution and the quartiles based upon all 229 controls.
[d] Based on the seroreactivity quartiles modelled as a continuous variable.

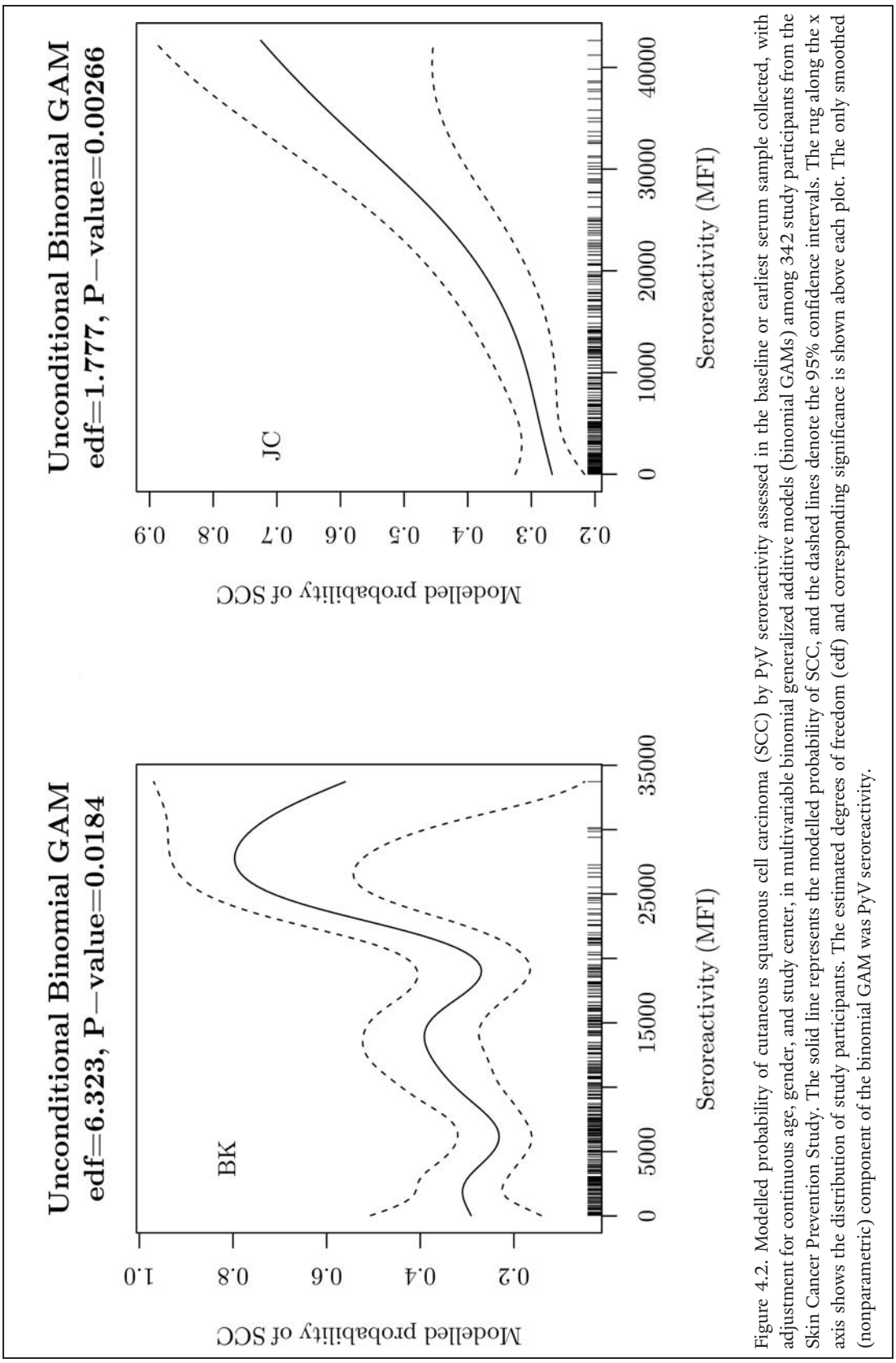



Figure 4.2. Modelled probability of cutaneous squamous cell carcinoma (SCC) by PyV seroreactivity assessed in the baseline or earliest serum sample collected, with adjustment for continuous age, gender, and study center, in multivariable binomial generalized additive models (binomial GAMs) among 342 study participants from the Skin Cancer Prevention Study. The solid line represents the modelled probability of SCC, and the dashed lines denote the 95% confidence intervals. The rug along the x axis shows the distribution of study participants. The estimated degrees of freedom (edf) and corresponding significance is shown above each plot. The only smoothed (nonparametric) component of the binomial GAM was PyV seroreactivity.

## 4.5 DISCUSSION

We conducted a nested case-control study to test the hypothesis that prediagnostic infection with human PyVs is associated with incident SCC in a population at high risk for this malignancy. Among participants with a history of KC, we found an increased risk of subsequent SCC associated with JC seropositivity, as well as with increasing quartiles of BK and JC seroreactivity, in serum samples collected prior to SCC diagnosis. SCC diagnosis was not associated with a change in BK or JC serostatus, and intraindividual seroreactivity remained consistent over time.

There is limited information on the longitudinal serostability of these viruses within individuals. Available studies on repeated measures of seroreactivity against PyVs have primarily been conducted among special populations such as organ transplant recipients,[316, 391] pregnant women,[34] and heavily immunosuppressed patients (e.g., HIV-infected individuals,[392, 393] or multiple sclerosis patients treated with natalizumab[377, 378, 379, 394, 395, 396]), with evidence of greater seroinstability over time, possible due to their condition or immunotherapy. A nested case-control study of non-Hodgkin's lymphoma found only 1 change in JC serostatus and 3 changes in BK serostatus over a period of 15 years out of 94 controls.[397] One published study aimed to investigate the serostability of PyVs in the general adult population: a longitudinal study of 458 individuals from Australia.[373] Over an 11 year follow-up period, BK seroprevalence was stable with only 2.5% displaying a change in serostatus over time.[373] However, JC serostatus was less stable during the same span of time, with 16% changing serostatus.[373] In our study, we found BK and JC seroreactivity to remain very stable within an individual, as nearly all participants who tested seropositive or seronegative in their first sample retained this status over the course of time. Thus, our findings along with prior work[373, 397] suggest a single measurement

of serum antibody level is a reasonable indicator of long-term antibody status for BK and JC in epidemiologic studies of adults.

Research exploring the effect of PyV serostatus prior to diagnosis on future SCC risk is limited, despite evidence of a causal role for MCV in MCC development.[12, 338] We found a positive association between JC seropositivity in prediagnostic serum samples and the development of a new primary SCC among those with least one previous KC. Our findings differ from those observed in a large Swedish study, where no association was found between SCC (OR=1.0, 95% CI: 0.8-1.4) and JC seropositivity with samples collected at least 1 month prior to diagnosis.[339] However, the Swedish study included JC as a specificity control, as they had the primary objective to investigate the association between HPV and SCC or BCC.[339] In a clinic-based case-control study conducted in Florida, USA, a positive but weak association for SCC was observed with JC seropositivity assessed at the time of diagnosis (OR=1.4, 95% CI: 0.9-2.2).[54] Our findings support this result, and further suggest a lack of bias from analysis of post-diagnosis samples. We also explored the risk of SCC in relation to the presence of prediagnostic antibodies against BK, but the high seroprevalence resulted in elevated but less stable estimates of SCC risk.

Infection with BK and JC has been categorized as 'possibly carcinogenic to humans' by the International Agency for Research in Cancer,[398] although the mechanism is not well understood. The early PyV protein, LTAg contains a retinoblastoma (Rb)-binding pocket necessary for transforming activity and cell proliferation,[3, 12, 15, 63] and also harbors a p53-binding domain that binds and inactivates the p53 protein to induce cell division.[3, 16, 63, 64] JC LTAg has been shown to directly interact with insulin receptor substrate-1 [399] and β-catenin,[400, 401] resulting in a nuclear translocation of proteins that may contribute to the process of malignant transformation through the dysregulation of homologous recombination-directed DNA repair[402] and the proto-

oncogene *c-MYC*.[401] BK and JC differ from other PyV types, as their late regions encode an agnoprotein which may exert its tumorigenic influence through cell cycle dysregulation, interference with DNA repair processes, and chromosome instability.[67, 68] Moreover, BK, JC, and MCV also encode a miRNA that downregulates LTAg expression, which may allow the virus to escape immune surveillance.[70]

A history of PyV infection assessed through antibody detection against a viral antigen implies that either viral action or confounding by an immune trait[403] related to both infection and SCC risk may be responsible for skin carcinogenesis. BK and JC are known to reactivate under conditions of immunosuppression (e.g., organ transplantation) due to impaired cellular immune responses and decreased immune surveillance.[7, 318, 354] Further, we previously found specific PyVs to be associated with slight changes in adaptive lymphocyte proportions among immunocompetent individuals using a bioinformatics approach.[341] As adaptive immunity is modulated by both genetics and exposure to infections or allergens, and plays a role in future cancer risk, research causally linking viruses and SCC is complicated by their shared association with immune dysfunction.[403]

### 4.5.1 Strengths and limitations of the study

We leveraged a unique trial with follow-up and serial serum samples to prospectively assess the relationship between SCC and PyV seroreactivity in a nested case-control study. Our prospective design reduced the possibility of reverse-causality with respect to the new occurrence of SCC – a concern of particular relevance for skin cancers, where the seroprevalence of HPV infections has been suspected to increase following diagnosis.[404] Our data suggest that PyV reactivation or increased susceptibility caused by the disease process are less likely to explain the observed associations. Nonetheless, the use of a high risk study sample comprised of patients with a history

of KCs in an intervention trial may limit the generalizability of our results. As all study participants had a history of KC, our study design does not exclude the possibility that the initial KC diagnosis prior to trial enrollment affected seroreactivity measurements.

Antibody response against each PyV type was used as a marker of PyV infection,[39, 59] measured using multiplex serology and a GST fusion protein-based capture immunoassay of recombinantly expressed VP1 capsid proteins. This has been shown a reliable technique to assess PyV seroreactivity and used as a marker of PyV infection in prior studies.[39, 53, 54, 59] We found antibody levels against BK and JC to be strongly correlated within individuals over time, and did not find evidence of cross-reactivity between BK and JC VP1 antibodies, suggesting our results to be virus-specific. A limitation of our study is the lack of measures for BK or JC virus in SCC tumors. Serum MFI values have been shown to correspond to the presence of viral DNA within tumor tissues, with MCV DNA-positive SCC tumors (but not MCV DNA-negative SCC tumors) having notably higher MFI values compared to controls;[54] yet far less is known about the other PyV types. While other studies have failed to detect DNA from either virus types in Bowen's disease,[364] BCC,[365] SCC,[54, 363, 365] or melanoma,[82, 363] BK DNA has been amplified from 76% of healthy skin tissues, and JC DNA from 16% of normal skin.[81] This suggests that these viruses can inhabit the skin even without being skin tropic.[3] In addition, BK (but not JC) early promoters displayed strong activity in experimental skin cell lines, which may indicate skin is suitable for polyomaviral propagation.[367] Another limitation is that a history of PyV infection inferred from antibody detection suggests either the virus itself or a confounding immunologic factor may be responsible for SCC development. Thus, our findings of an association between BK or JC and SCC could represent the predisposing role of an immune phenotype, rather than the carcinogenic action of the virus.

### 4.5.2 Conclusions

We found evidence for an association between prediagnostic seroreactivity to BK and JC, and subsequent risk of SCC development in a USA adult population at high risk for KC. Further, our data suggest that seroreactivity to BK and JC is relatively stable over a period of up to 6 years, and following an SCC diagnosis.

# Chapter 5

# Concluding remarks

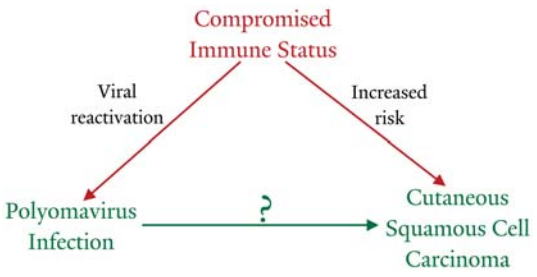



Figure 5.1. Does immune status confound the relationship between polyomavirus infection and risk of cutaneous squamous cell carcinoma?

In this thesis, we were interested in the seroepidemiology of human polyomaviruses (PyV) in a United States population, the association between human PyVs and incidence of cutaneous squamous cell carcinoma (SCC) in the New Hampshire Skin Cancer Study, and the prospective relation of human PyVs and risk of SCC in the Skin Cancer Prevention Study. We found that a history of PyV infection is common in the United States, and varies by sociodemographic (i.e., age, gender, smoking status, glucocorticoid use) and biologic (i.e., immune profiles) characteristics. We also learned that John Cunningham virus (JC) seropositivity assessed following SCC diagnosis increased the odds of incident SCC, with a trend in increasing risk with increasing quartiles of seroreactivity. Reverse causality was unlikely to account for this association, as we further observed both BK and JC seroreactivity to be stable longitudinally and following an SCC diagnosis. JC seropositivity prior to diagnosis was associated with an elevated risk of SCC, and a trend in increasing risk with increasing quartiles of BK and JC seroreactivity was observed amongst those with a prior keratinocyte cancer. However, these thesis results should be interpreted cautiously as potential confounding by immune competence remains (Figure 5.1), and a noteworthy limitation of this research is the use of serologic evidence of PyV infection rather than assaying for PyV DNA in the tissue of interest.

## 5.1 FUTURE DIRECTIONS

All known human PyVs can act as oncoviruses because they encode the oncoprotein, T antigen; yet the only human PyV that has demonstrated oncogenicity in humans is Merkel cell polyomavirus (MCV), which is a causal factor for Merkel cell carcinoma (MCC). Apart from the

role of MCV in MCC, none of the other PyVs have yet been implicated in human carcinogenesis. In this thesis, we found high seroprevalence rates for most PyVs among adults in the US, implying that PyVs circulate widely within the general population – including the pathogenic PyVs. Despite the experimental evidence for a potential etiologic role in SCC, it is difficult to establish a causal relation between the virus and the oncogenic events on a serological level due to the ubiquitous prevalence of most PyVs. Our findings do not exclude the possibility that JC may be involved in the disease. However, there has been no evidence in the literature that the SCC tumors contain integrated JC virus genomes. Paraffin-embedded tissue samples were also collected from a subset of study participants in the New Hampshire Skin Cancer Study[405] and could be tested for PyV DNA using multiplex PCR.[54, 362] However, the presence of episomal PyV DNA is insufficient to support a causal role for these viruses in the pathogenesis of a cancer, because the tumor itself might permit high levels of viral replication. MCV, HPyV6, HPyV7, and TSV have been found to show high tropism for the skin but were not associated with SCC in this thesis. The role of PyVs in human cancer (if any) is an important area of future research requiring the screening of many different types of cancer, such as bladder cancer.[308] It is important to understand that PyVs may simply not be involved in the etiology of skin cancer regardless of the findings of this thesis.

The precise mechanisms of PyV persistence and reactivation remains to be understood, and the cell types that serve as viral reservoirs must be discovered. Defining the latency state, such as a low level of continuous lytic infection, would assist in understanding the host immunological response to PyV infection and its' association with disease. It is possible infection with and persistence of normally benign PyVs confers some advantage to the host. The increasing number

of human and animal PyVs suggests that there are many more to be discovered and raises key questions regarding their tropism, spread, and disease-causing potential.

An important characteristic of human PyVs (especially BK and JC viruses) is that while they initially cause asymptomatic infections, they can cause disease in immunocompromised patients. Insight into reactivation of latent viruses or of normally 'benign' persistent viruses in immunocompromised patients has become increasingly important in the field of organ transplantations and for HIV infected patients. Whether active PyV replication occurs in immunocompromised patients such as organ transplant recipients, AIDS patients, and patients with autoimmune disease remains to be thoroughly examined. In general, more research is needed to understand the role of immunosurveillance in controlling PyVs by utilising special populations.

## 5.2 IMPLICATIONS

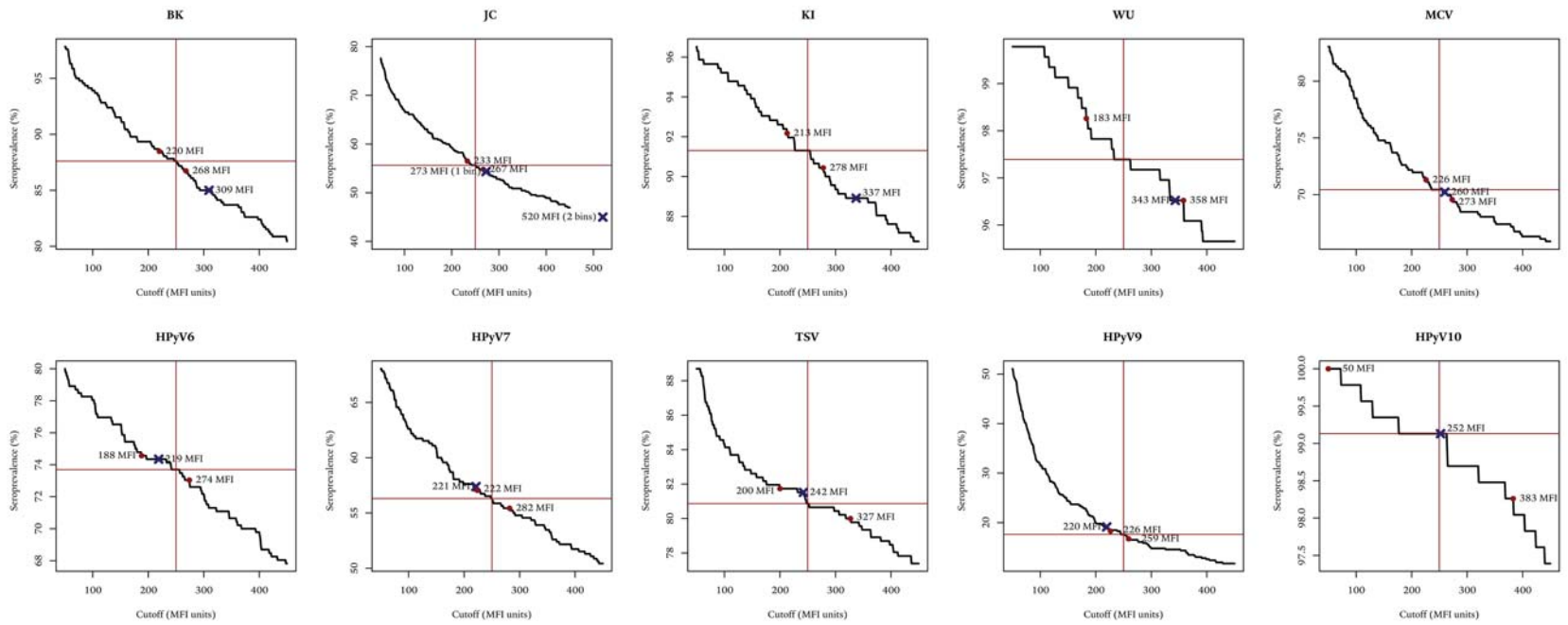
Given the increasing use of immunosuppressive therapy after organ transplantation, and for autoimmune diseases such as multiple sclerosis and rheumatoid arthritis, the frequency of diseases resulting from PyVs and risk factors for infection or reactivation may become more apparent over time. A market for effective therapeutics and vaccines that inhibit PyV replication, target the oncogenic activity of T antigens, boost the immune response against certain PyVs, or prevent infection altogether may emerge – especially when considering the causative role of JC in the life-threatening condition of progressive multifocal leukoencephalopathy and granule cell neuronopathy in immunosuppressed patients.[406] The plausibility of antiviral therapy and vaccine development should also be considered. Repeatedly, the medical world will be challenged with emerging human viruses with novel disease associations, and therapies will need to be developed.

# Appendices

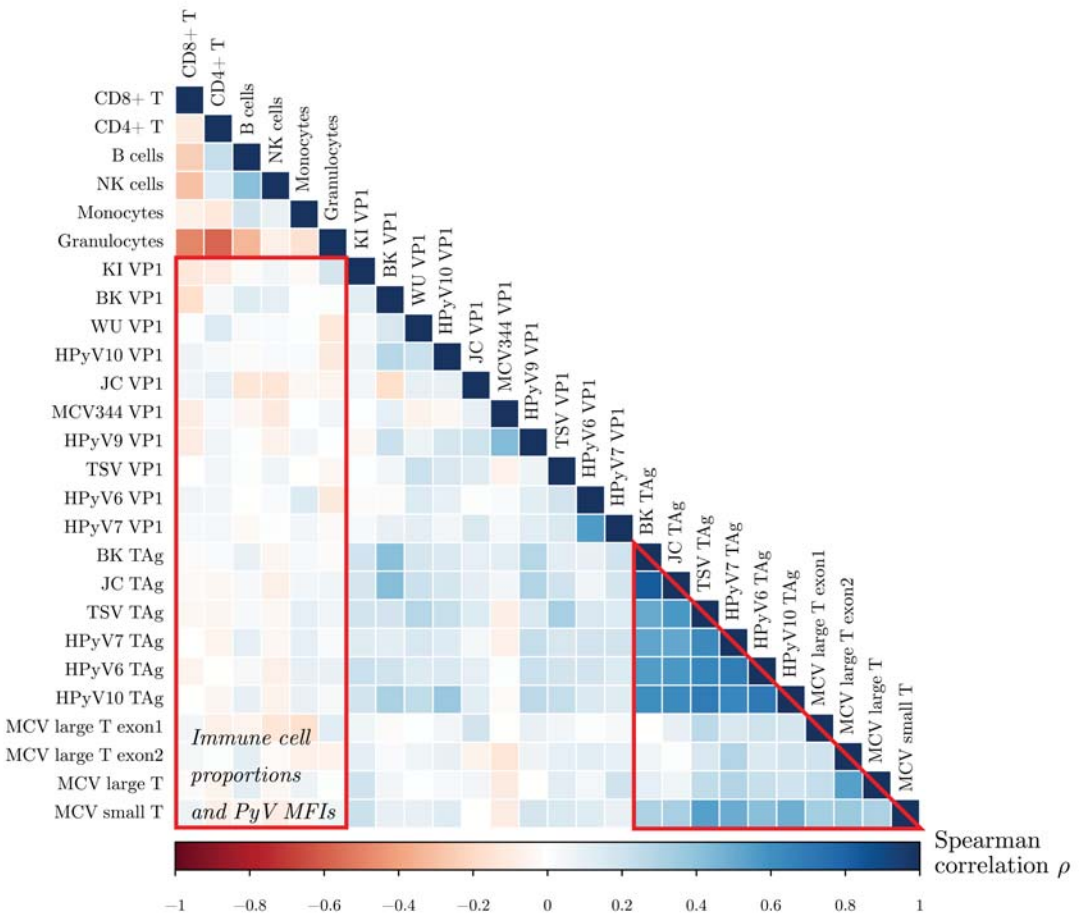
Supplemental Table 2.1. Distribution of human polyomavirus (PyV) seropositivity among 194 study participants by sociodemographic factors. [a]

| Variable | Total, No. (%) [d] | No. PyV types seropositive, mean (SD) [e] | Seropositive, No. (%) [b,c] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | BK | JC | KI | WU | MCV | HPyV6 | HPyV7 | TSV | HPyV9 | HPyV10 |
| Overall | 194 (100) | 7.3 (1.4) | 164 (84.5) | 105 (54.1) | 177 (91.2) | 189 (97.4) | 130 (67.0) | 144 (74.2) | 120 (61.8) | 163 (84.0) | 38 (19.6) | 192 (99.0) |
| Gender | | | | | | | | | | | | |
| Male | 121 (100) | 7.4 (1.6) | 101 (83.5) | 69 (57.0) | 111 (91.7) | 118 (97.5) | 82 (67.8) | 91 (75.2) | 79 (65.3) | 107 (88.4) | 24 (19.8) | 120 (99.2) |
| Female | 73 (100) | 7.1 (1.2) | 63 (86.3) | 36 (49.3) | 66 (90.4) | 71 (97.3) | 48 (65.7) | 53 (72.6) | 41 (56.2) | 56 (76.7) | 14 (19.2) | 72 (98.6) |
| Age (years) | | | | | | | | | | | | |
| <45 | 23 (100) | 6.5 (1.3) | 22 (95.6) | 8 (34.8) | 19 (82.6) | 23 (100) | 14 (60.9) | 17 (73.9) | 8 (34.8) | 17 (73.9) | 2 (8.7) | 23 (100) |
| 45-54 | 28 (100) | 7.4 (1.7) | 26 (92.8) | 15 (53.6) | 24 (85.7) | 27 (96.4) | 19 (67.9) | 23 (82.1) | 17 (60.7) | 23 (82.1) | 6 (21.4) | 27 (96.4) |
| 55-64 | 45 (100) | 7.4 (1.3) | 38 (84.4) | 27 (60.0) | 43 (85.6) | 44 (97.8) | 34 (75.6) | 31 (68.9) | 25 (55.6) | 40 (88.9) | 8 (17.8) | 44 (97.8) |
| 65-69 | 52 (100) | 7.3 (1.5) | 42 (80.8) | 32 (61.5) | 47 (90.4) | 50 (96.1) | 30 (57.7) | 37 (71.2) | 38 (73.1) | 44 (84.6) | 9 (17.3) | 52 (100) |
| 70-75 | 46 (100) | 7.5 (1.4) | 36 (78.3) | 23 (50.0) | 44 (95.6) | 45 (97.8) | 33 (71.7) | 36 (78.3) | 32 (69.6) | 39 (84.8) | 13 (28.3) | 46 (100) |
| Education | | | | | | | | | | | | |
| Elementary to high or technical school | 98 (100) | 7.3 (1.4) | 83 (84.7) | 53 (54.1) | 94 (95.9) | 95 (96.9) | 64 (65.3) | 69 (70.4) | 57 (58.2) | 87 (88.8) | 18 (18.4) | 97 (99.0) |
| Any college | 59 (100) | 7.1 (1.4) | 50 (84.7) | 28 (47.5) | 52 (88.1) | 57 (96.6) | 40 (67.8) | 44 (74.6) | 37 (62.7) | 46 (78.0) | 7 (11.9) | 59 (100) |
| Graduate or professional school | 37 (100) | 7.7 (1.6) | 31 (83.8) | 24 (64.9) | 31 (83.8) | 37 (100) | 26 (70.2) | 31 (83.8) | 26 (70.3) | 30 (81.1) | 13 (35.1) | 36 (97.3) |
| Smoking status | | | | | | | | | | | | |
| Never | 64 (100) | 7.2 (1.3) | 56 (87.5) | 32 (50.0) | 57 (89.1) | 63 (98.4) | 49 (76.6) | 43 (67.2) | 37 (57.8) | 57 (89.1) | 6 (9.4) | 62 (96.9) |
| Former | 97 (100) | 7.4 (1.5) | 82 (84.5) | 53 (54.6) | 90 (92.8) | 93 (95.9) | 64 (66.0) | 76 (78.4) | 65 (67.0) | 77 (79.4) | 25 (25.8) | 97 (100) |
| Current | 33 (100) | 7.2 (1.4) | 26 (78.8) | 20 (60.6) | 30 (90.9) | 33 (100) | 17 (51.5) | 25 (75.8) | 18 (54.5) | 29 (87.9) | 7 (21.2) | 33 (100) |
| Skin sun sensitivity [f] | | | | | | | | | | | | |
| Burn | 70 (100) | 7.0 (1.5) | 57 (81.4) | 37 (52.8) | 60 (85.7) | 68 (97.1) | 44 (62.8) | 48 (68.6) | 39 (55.7) | 59 (84.3) | 9 (12.8) | 69 (98.6) |
| Tan | 123 (100) | 7.5 (1.4) | 106 (86.2) | 68 (52.3) | 116 (94.3) | 120 (97.6) | 85 (69.1) | 95 (77.2) | 80 (65.0) | 103 (83.7) | 29 (23.6) | 122 (99.2) |
| No. of lifetime painful sunburns | | | | | | | | | | | | |
| 0-1 | 100 (100) | 7.4 (1.4) | 85 (85.0) | 56 (56.0) | 89 (89.0) | 98 (98.0) | 70 (70.0) | 77 (77.0) | 61 (61.0) | 80 (80.0) | 23 (23.0) | 99 (99.0) |
| 2+ | 93 (100) | 7.3 (1.5) | 78 (83.9) | 48 (51.6) | 87 (93.5) | 90 (96.8) | 60 (64.5) | 67 (72.0) | 58 (62.4) | 82 (88.2) | 14 (15.0) | 92 (98.9) |
| Glucocorticoid use [g] | | | | | | | | | | | | |
| Yes | 22 (100) | 7.4 (1.4) | 19 (86.3) | 12 (54.5) | 21 (95.4) | 21 (95.4) | 13 (59.1) | 16 (72.7) | 18 (81.8) | 18 (81.8) | 3 (13.6) | 22 (100) |
| No | 168 (100) | 7.3 (1.4) | 142 (84.5) | 92 (54.8) | 152 (90.5) | 165 (98.2) | 114 (67.9) | 126 (75.0) | 101 (60.1) | 141 (83.9) | 34 (20.2) | 166 (98.8) |
| Transplant recipient | | | | | | | | | | | | |
| Yes | 0 | NA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| No | 193 (100) | 7.3 (1.4) | 164 (85.0) | 105 (54.4) | 176 (91.2) | 188 (97.4) | 129 (66.8) | 143 (74.1) | 120 (62.2) | 162 (83.9) | 38 (19.7) | 191 (99.0) |

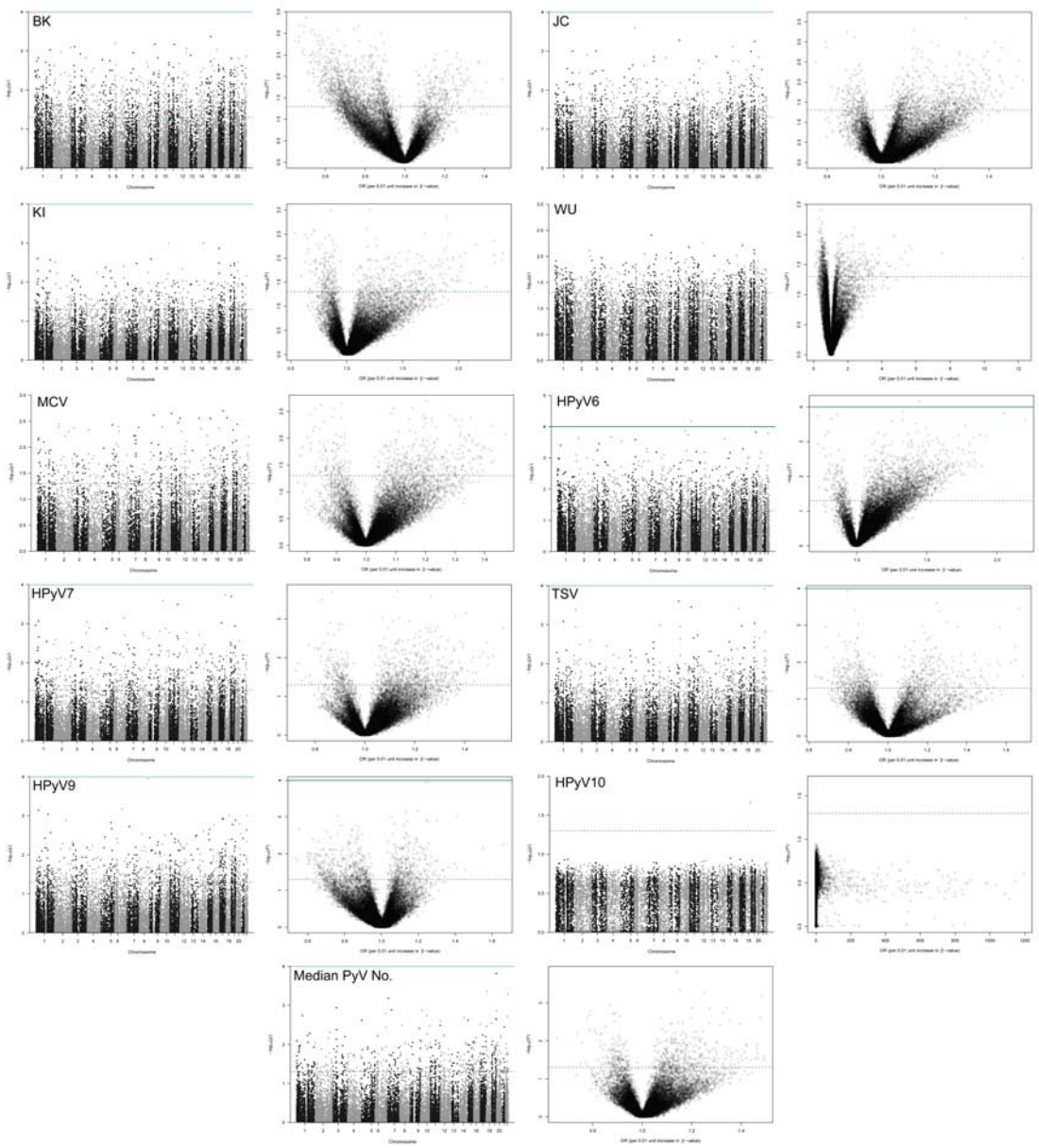
[a] Numbers may not sum to the overall total due to missing data.
[b] Percentages indicate the proportion of healthy adults who are PyV seropositive versus PyV seronegative in each given strata.
[c] PyV infection was determined using seropositivity for the VP1 protein.
[d] All participants were seropositive for at least one of the ten PyVs assayed.
[e] Standard deviation referred to as "SD".
[f] Sun sensitivity was defined as the reaction to 1 hour of sun exposure the first time in the summer. "Burn" refers to those who experienced severe sunburn with blistering or a painful sunburn and then peeling. "Tan" refers to those who experienced mild sunburn and then some tanning or tanned without sunburn.
[g] Glucocorticoid use was defined as having used oral steroid or corticosteroid medications (for example, cortisone or prednisone) for one month or longer.




Supplemental Figure 2.1. Robustness of human polyomavirus (PyV) seroprevalence estimates. The cut points were varied from 50 to 450 median fluorescence intensity (MFI) units (x axis), and the resulting seroprevalences were calculated using the new cutoffs (y axis). The red cross hairs show the seroprevalences for each PyV VP1 using the recommended cutoff=250 MFI units. The two red dots show the range of MFI units that would result in seroprevalences ±1% of the seroprevalences calculated using the recommended cutoff. The blue crosses denote the cut points calculated using a frequency distribution analysis as described in van der Meijden *et al.*, 2013, with two JC cutoffs calculated using 1 bin versus 2 bins to illustrate the large difference in seroprevalence estimates depending upon binned samples included.




Supplemental Figure 2.2. Spearman rank correlation coefficients, ρ, between immune cell proportions and median fluorescence intensity (MFI) values for each human polyomavirus (PyV) VP1 or T antigen (TAg). Not all PyV TAgs were assayed, and MCV large TAg was assayed using the entire protein as well as with two fragments (exon 1 and exon 2) of the full length large TAg. The red square emphasizes the correlations between immune cell proportions and MFI values for each PyV, while the red triangle emphasizes the correlations between PyV TAgs.




Supplemental Figure 2.3. Manhattan and volcano plots for the results of an epigenome wide association study (EWAS) following locus-by-locus examination of differentially methylated CpG sites based on polyomavirus (PyV) seroreactivity, as well as seropositivity for </≥ 7 PyVs. Included in the analysis were 20,046 CpG sites used to predict seropositivity against 10 different human PyV or seropositivity for </≥ 7 PyVs in 194 study subjects. CpG loci located on the sex chromosomes, assayed with probes containing SNPs, assayed with cross-reactive probes, or assayed with repeated probes were removed along with poor-performing loci. The 2 plots present *P*-values and association estimates after adjusting for age, sex, and immune cell proportions. The *P*-values presented were derived from logistic models using logit-transformed β-values, and the paired odds ratios (OR) from models using untransformed β-values to allow for biologic interpretability. The dotted green line represents an α=0.05 level of significance, while the solid green line represents an α=0.0001 level of significance. No significant associations were significant at the FDR level.

## CHAPTER 3 SUPPLEMENTAL MATERIALS

Supplemental Table 3.1. Distribution of polyomavirus (PyV) VP1 and T antigen (TAg) seropositivities and seronegativities amongst 713 study participants from the New Hampshire Skin Cancer Study. Data on whether an individual was seroreactive against the PyV protein TAg (a PyV oncoprotein) was collected. Standard cut points for seropositivity were chosen for the MFI of each PyV tested. The cutoffs were defined arbitrarily from visual inspection of frequency distribution curves (percentile plots) of all sera tested, and stringent criteria were chosen to increase specificity. The cutoff value for TAg was 400 MFI units for BK, JC, MCV large TAg, HPyV6, HPyV7, TSV, and HPyV10; the cutoff value for MCV small TAg was 50 MFI units. Generally, VP1 seropositivity is required to be considered to have a history of PyV infection. However, it can be seen that those who are VP1 seropositive are rarely also TAg seropositive.

| **BK** | TAg negative | TAg positive | Concordance (%) |
|---|---|---|---|
| VP1 negative | 89 | 4 | 18.2 |
| VP1 positive | 579 | 41 | |
| **JC** | TAg negative | TAg positive | |
| VP1 negative | 257 | 33 | 42.6 |
| VP1 positive | 376 | 47 | |
| **MCV** | Small TAg negative | Small TAg positive | |
| VP1 negative | 214 | 1 | 30.7 |
| VP1 positive | 493 | 5 | |
| **MCV** | Large T Exon 1 negative | Large T Exon 1 positive | |
| VP1 negative | 215 | 0 | 30.3 |
| VP1 positive | 497 | 1 | |
| **MCV** | Large T Exon 2 negative | Large T Exon 2 positive | |
| VP1 negative | 213 | 2 | 30.2 |
| VP1 positive | 494 | 4 | |
| **MCV** | Large T negative | Large T positive | |
| VP1 negative | 214 | 1 | 30.3 |
| VP1 positive | 496 | 2 | |
| **TSV** | TAg negative | TAg positive | |
| VP1 negative | 136 | 1 | 19.5 |
| VP1 positive | 573 | 3 | |
| **HPyV6** | TAg negative | TAg positive | |
| VP1 negative | 182 | 2 | 27.5 |
| VP1 positive | 515 | 14 | |
| **HPyV7** | TAg negative | TAg positive | |
| VP1 negative | 294 | 4 | 43.3 |
| VP1 positive | 400 | 15 | |
| **HPyV10** | TAg negative | TAg positive | |
| VP1 negative | 7 | 0 | 3.1 |
| VP1 positive | 691 | 15 | |

Supplemental Table 3.2. Odds ratios (95% confidence intervals) for cutaneous squamous cell carcinoma (SCC) by quartiles of polyomavirus (PyV) seroreactivity among 713 study participants from the New Hampshire Skin Cancer Study.

| PyV seroreactivity (MFI units) | Controls (n=460), No. (%) | SCC Cases (n=253) | |
|---|---|---|---|
| | | No. (%) | OR (95% CI)* |
| **BK** | | | |
| Quartile 1 | 115 (25.0) | 64 (25.3) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 64 (25.3) | 1.01 (0.65-1.57) |
| Quartile 3 | 115 (25.0) | 63 (24.9) | 0.99 (0.64-1.55) |
| Quartile 4 | 115 (25.0) | 62 (24.5) | 1.15 (0.74-1.81) |
| *P* for trend | | | 0.58 |
| **KI** | | | |
| Quartile 1 | 115 (25.0) | 52 (20.6) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 83 (32.8) | 1.62 (1.04-2.53) |
| Quartile 3 | 115 (25.0) | 61 (24.1) | 1.27 (0.80-2.01) |
| Quartile 4 | 115 (25.0) | 57 (22.5) | 1.08 (0.67-1.72) |
| *P* for trend | | | 0.89 |
| **WU** | | | |
| Quartile 1 | 115 (25.0) | 74 (29.2) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 49 (19.4) | 0.68 (0.43-1.07) |
| Quartile 3 | 115 (25.0) | 68 (26.9) | 0.95 (0.62-1.47) |
| Quartile 4 | 115 (25.0) | 62 (24.5) | 0.85 (0.55-1.32) |
| *P* for trend | | | 0.78 |
| **MCV** | | | |
| Quartile 1 | 115 (25.0) | 66 (26.1) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 77 (30.4) | 1.08 (0.70-1.66) |
| Quartile 3 | 115 (25.0) | 41 (16.2) | 0.57 (0.35-0.93) |
| Quartile 4 | 115 (25.0) | 69 (27.3) | 0.97 (0.63-1.51) |
| *P* for trend | | | 0.37 |
| **HPyV6** | | | |
| Quartile 1 | 115 (25.0) | 54 (21.3) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 64 (25.3) | 1.13 (0.72-1.79) |
| Quartile 3 | 115 (25.0) | 65 (25.7) | 1.20 (0.76-1.89) |
| Quartile 4 | 115 (25.0) | 70 (27.7) | 1.22 (0.78-1.92) |
| *P* for trend | | | 0.37 |
| **HPyV7** | | | |
| Quartile 1 | 115 (25.0) | 51 (20.2) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 62 (24.5) | 1.18 (0.74-1.88) |
| Quartile 3 | 115 (25.0) | 57 (22.5) | 0.90 (0.56-1.44) |
| Quartile 4 | 115 (25.0) | 83 (32.8) | 1.41 (0.90-2.21) |
| *P* for trend | | | 0.25 |
| **TSV** | | | |
| Quartile 1 | 115 (25.0) | 60 (23.7) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 66 (26.1) | 0.98 (0.62-1.54) |
| Quartile 3 | 115 (25.0) | 70 (27.7) | 1.09 (0.70-1.72) |
| Quartile 4 | 115 (25.0) | 57 (22.5) | 1.02 (0.64-1.62) |
| *P* for trend | | | 0.82 |
| **HPyV9** | | | |
| Quartile 1[a] | 114 (24.8) | 53 (20.9) | 1.00 (referent) |
| Quartile 2 | 113 (24.6) | 72 (28.4) | 1.31 (0.83-2.05) |
| Quartile 3 | 118 (25.6) | 48 (19.0) | 0.85 (0.52-1.37) |
| Quartile 4 | 115 (25.0) | 80 (31.6) | 1.41 (0.90-2.21) |
| *P* for trend | | | 0.35 |
| **HPyV10** | | | |
| Quartile 1 | 115 (25.0) | 66 (26.1) | 1.00 (referent) |
| Quartile 2 | 115 (25.0) | 72 (28.4) | 1.14 (0.74-1.76) |
| Quartile 3 | 115 (25.0) | 50 (19.8) | 0.82 (0.52-1.31) |
| Quartile 4 | 115 (25.0) | 65 (25.7) | 1.16 (0.75-1.82) |
| *P* for trend | | | 0.84 |

* Adjusted for age group and gender. OR=odds ratio, CI=confidence interval.

[a] Controls may not be evenly distributed within quartiles due to uneven data distribution.

Supplemental Table 3.3. Odds ratios (95% confidence intervals) for cutaneous squamous cell carcinoma (SCC) by seropositivity for all cutaneous polyomaviruses (PyV), and by number of cutaneous PyV types seropositive, among 713 study participants from the New Hampshire Skin Cancer Study.

| PyV seroreactivity | Controls (n=460), No. (%) | SCC Cases (n=253) | |
|---|---|---|---|
| | | No. (%) | OR (95% CI)* |
| **Cutaneous**‡ | | | |
| Seronegative | 313 (68.0) | 168 (66.4) | 1.00 (referent) |
| Seropositive | 147 (32.0) | 85 (33.6) | 0.94 (0.67-1.32) |
| Continuous | | | 0.98 (0.85-1.14) |
| *P* | | | 0.83 |

* Adjusted for age group and gender. OR=odds ratio, CI=confidence interval.

‡ Cutaneous PyVs consisted of MCV, HPyV6, HPyV7, and TSV. Those who were seropositive for all 4 cutaneous PyVs were compared to those seropositive for ≤3 cutaneous PyVs.

Supplemental Table 3.4. Characteristics of cutaneous squamous cell carcinoma (SCC) case and control study participants by sociodemographic and skin cancer risk factors, as well as by single nucleotide polymorphisms (SNP) or haplotype (n=713).[a]

| Variable | SCC Cases (n=253), No. (%) | Controls (n=460), No. (%) | HLA-C (n=615) | | | HLA-C -35 kb (n=652) | | | NKG2Da (n=654) | | | NKG2Db (n=638) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | C1 (n=253, 35.5) | C1C2 (n=273, 38.3) | C2 (n=89, 12.5) | TT (n=236, 33.1) | CT (n=320, 44.9) | CC (n=96, 13.5) | CC (n=434, 60.9) | CT (n=194, 27.2) | TT (n=26, 3.6) | CC (n=340, 47.7) | CG (n=235, 33.0) | GG (n=63, 8.8) |
| Gender | | | | | | | | | | | | | | |
| Male | 168 (66.4) | 280 (60.9) | 168 (37.5) | 168 (37.5) | 52 (11.6) | 163 (36.4) | 193 (43.1) | 56 (12.5) | 269 (60.0) | 125 (27.9) | 18 (4.0) | 218 (48.7) | 145 (32.4) | 42 (9.4) |
| Female | 85 (33.6) | 180 (39.1) | 85 (32.1) | 105 (39.6) | 37 (14.0) | 73 (27.5) | 127 (47.9) | 40 (15.1) | 165 (62.3) | 69 (26.0) | 8 (3.0) | 122 (46.0) | 90 (33.9) | 21 (4.5) |
| Median age, SD (years) | 68 (8.0) | 65 (10.7)*** | 66 (9.7) | 66 (10.0) | 66 (11.1) | 67 (9.9) | 66 (9.7) | 65 (10.6) | 66 (10.1) | 66 (9.9) | 63.5 (10.3) | 66 (9.9) | 66 (10.3) | 66 (9.4) |
| Education | | | | | | | | | | | | | | |
| Elementary to high or technical school | 111 (43.9) | 227 (49.3) | 117 (34.6) | 130 (38.5) | 40 (11.8) | 116 (34.3) | 148 (43.8) | 46 (13.6) | 201 (59.5) | 100 (29.6) | 11 (3.2) | 154 (45.6) | 118 (34.9) | 27 (8.0) |
| Any college | 80 (31.6) | 144 (31.1) | 76 (33.9) | 93 (41.5) | 26 (11.6) | 71 (31.7) | 102 (45.5) | 32 (14.3) | 140 (62.5) | 56 (25.0) | 10 (4.5) | 116 (51.8) | 63 (28.1) | 22 (9.8) |
| Graduate or professional school | 61 (24.1) | 89 (19.3) | 60 (40.0) | 49 (32.7) | 23 (15.3) | 48 (32.0) | 70 (46.7) | 18 (12.0) | 92 (61.3) | 38 (25.3) | 5 (3.3) | 69 (46.0) | 54 (36.0) | 14 (9.3) |
| Smoking status [b] | | | | | | | | | | | | | | |
| Never | 79 (31.2) | 146 (31.7) | 68 (30.2) | 93 (41.3) | 36 (16.0) | 74 (32.9) | 110 (48.9) | 23 (10.2) | 135 (60.0) | 58 (25.8) | 12 (5.3) | 102 (45.3) | 75 (33.3) | 26 (11.6)* |
| Former | 133 (52.6) | 230 (50.0) | 134 (36.9) | 134 (36.9) | 43 (11.8) | 115 (31.7) | 161 (44.3) | 58 (16.0) | 222 (61.1) | 100 (27.5) | 12 (3.3) | 179 (49.3) | 111 (30.6) | 35 (9.6) |
| Current | 40 (15.8) | 84 (18.3) | 51 (41.1) | 45 (36.2) | 10 (8.1) | 46 (37.1) | 49 (39.5) | 15 (12.1) | 76 (61.3) | 36 (29.0) | 2 (1.6) | 58 (46.8) | 49 (39.5) | 2 (1.6) |
| Skin color | | | | | | | | | | | | | | |
| Light | 213 (84.2) | 279 (60.6)*** | 190 (38.6) | 187 (38.0) | 49 (9.9)*** | 168 (34.1) | 225 (45.7) | 58 (11.8) | 300 (61.0) | 132 (26.8) | 19 (3.9) | 232 (47.1) | 170 (34.6) | 42 (8.5) |
| Medium | 38 (15.0) | 180 (39.1) | 63 (28.9) | 86 (39.4) | 38 (17.4) | 67 (30.7) | 94 (43.1) | 37 (17.0) | 132 (60.6) | 61 (28.0) | 7 (3.2) | 106 (48.6) | 64 (29.4) | 21 (9.6) |
| Skin sun sensitivity [c] | | | | | | | | | | | | | | |
| Severe sunburn with blistering | 22 (8.7) | 28 (6.1)*** | 17 (34.0) | 18 (36.0) | 6 (12.0) | 15 (30.0) | 23 (46.0) | 7 (14.0) | 30 (60.0) | 12 (24.0) | 3 (6.0) | 24 (48.0) | 16 (32.0) | 3 (6.0) |
| Painful sunburn and then peeling | 93 (36.8) | 116 (25.2) | 73 (34.9) | 88 (42.1) | 22 (10.5) | 77 (36.8) | 98 (46.9) | 21 (10.0) | 122 (58.4) | 66 (31.6) | 7 (3.3) | 102 (48.8) | 70 (33.5) | 19 (9.1) |
| Mild sunburn with some tanning | 111 (43.9) | 234 (50.9) | 126 (36.5) | 121 (35.1) | 52 (15.1) | 106 (30.7) | 153 (44.3) | 53 (15.4) | 214 (62.0) | 85 (24.6) | 11 (3.2) | 164 (47.5) | 115 (33.3) | 30 (8.7) |
| Tan without sunburn | 25 (9.9) | 80 (17.4) | 36 (34.3) | 44 (41.9) | 9 (8.6) | 36 (34.3) | 45 (42.9) | 15 (14.3) | 65 (61.9) | 30 (28.6) | 5 (4.8) | 48 (45.7) | 34 (32.4) | 10 (9.5) |
| No. of lifetime painful sunburns [d] | | | | | | | | | | | | | | |
| 0 | 66 (26.1) | 147 (32.0)*** | 75 (35.2) | 89 (41.8) | 26 (12.2) | 77 (36.2) | 96 (45.1) | 27 (12.7)* | 128 (60.1) | 60 (28.2) | 13 (6.1) | 92 (43.2) | 79 (37.1) | 22 (10.3) |
| 1-2 | 49 (19.4) | 134 (29.1) | 66 (36.1) | 66 (36.1) | 22 (12.0) | 73 (39.9) | 67 (36.6) | 22 (12.0) | 108 (59.0) | 55 (30.0) | 3 (1.6) | 87 (47.5) | 64 (35.0) | 11 (6.0) |
| 3+ | 136 (53.8) | 174 (37.8) | 108 (34.8) | 115 (37.1) | 41 (13.2) | 82 (26.4) | 154 (49.7) | 47 (15.2) | 194 (62.6) | 76 (24.5) | 10 (3.2) | 158 (51.0) | 89 (28.7) | 29 (9.3) |
| Glucocorticoid use [e] | | | | | | | | | | | | | | |
| Yes | 33 (13.0) | 39 (8.5) | 23 (31.9) | 30 (41.7) | 6 (8.3) | 19 (26.4) | 34 (47.2) | 10 (13.9) | 44 (61.1) | 16 (22.2) | 4 (5.6) | 35 (48.6) | 17 (23.6) | 12 (16.7)* |
| No | 211 (83.4) | 415 (90.2) | 225 (35.9) | 237 (37.8) | 81 (12.9) | 206 (32.9) | 285 (45.5) | 83 (13.3) | 383 (61.2) | 174 (27.8) | 21 (3.3) | 300 (47.9) | 211 (33.7) | 50 (8.0) |
| Transplant recipient | | | | | | | | | | | | | | |
| Yes | 6 (2.4) | 1 (0.2)** | 1 (14.3) | 3 (42.8) | 2 (28.6) | 4 (57.1) | 1 (14.3) | 2 (28.6) | 3 (42.8) | 1 (14.3) | 1 (14.3) | 2 (28.6) | 2 (28.6) | 1 (14.3) |
| No | 246 (97.2) | 458 (88.6) | 251 (35.6) | 269 (38.2) | 87 (12.4) | 230 (32.7) | 319 (45.3) | 94 (13.4) | 429 (60.9) | 193 (27.4) | 25 (3.6) | 337 (47.9) | 232 (32.9) | 62 (8.8) |
| **SNP Genotype** | | | | | | | | | | | | | | |
| HLA-C | | | | | | | | | | | | | | |
| C1 | 98 (38.7) | 155 (33.7) | | | | | | | | | | | | |
| C1C2 | 99 (39.1) | 174 (37.8) | | | | | | | | | | | | |
| C2 | 28 (11.1) | 61 (13.3) | | | | | | | | | | | | |
| HLA-C -35 kb | | | | | | | | | | | | | | |
| TT | 86 (34.0) | 150 (32.6) | | | | | | | | | | | | |
| CT | 118 (46.6) | 202 (43.9) | | | | | | | | | | | | |
| CC | 37 (14.6) | 59 (12.8) | | | | | | | | | | | | |
| NKG2Da | | | | | | | | | | | | | | |
| CC | 164 (64.8) | 270 (58.6) | | | | | | | | | | | | |
| CT | 64 (25.3) | 130 (28.3) | | | | | | | | | | | | |
| TT | 8 (3.2) | 18 (3.9) | | | | | | | | | | | | |
| NKG2Db | | | | | | | | | | | | | | |
| CC | 130 (51.4) | 210 (45.6) | | | | | | | | | | | | |
| CG | 85 (33.6) | 150 (32.6) | | | | | | | | | | | | |
| GG | 21 (8.3) | 42 (9.1) | | | | | | | | | | | | |

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. $P$values obtained from $X^2$, Fisher's exact, or Wilcoxon rank sum test (as appropriate) comparing sociodemographic and skin cancer risk factors between SCC cases and controls, comparing SNP genotype between cases and controls, and comparing sociodemographic skin cancer risk factory by SNP genotype.

[a] Numbers may not sum to the overall total due to missing data. They were excluded from complete-case analyses.

[b] Cigarette smoking status at 1 year prior to the reference or diagnosis date

[c] Sun sensitivity was defined as the skin reaction to 1 hour of sun exposure the first time in the summer.

[d] Sunburns that caused pain for 2 or more days

[e] Glucocorticoid use was defined as having used oral steroid or corticosteroid medications (e.g., cortisone or prednisone) for one month or longer.

Supplemental Table 3.5. Distribution of human polyomavirus (PyV) seropositivity among 460 control participants by single nucleotide polymorphism (SNP) or haplotype.[a]

| SNP Genotype | Total, No. (%) [d] | No. PyV seropositive, mean (SD) [e] | Seropositive, No. (%) [b,c] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | BK (n=403, 87.6) | JC (n=256, 55.6) | KI (n=420, 91.3) | WU (n=448, 97.4) | MCV (n=324, 70.4) | HPyV6 (n=339, 73.7) | HPyV7 (n=259, 56.3) | TSV (n=372, 80.9) | HPyV9 (n=81, 17.6) | HPyV10 (n=456, 99.1) |
| **HLA-C** | | | | | | | | | | | | |
| C1 | 155 (100) | 7.3 (1.5) | 133 (85.8) | 90 (58.1) | 141 (91.0) | 149 (96.1) | 102 (65.8) | 111 (71.6) | 93 (60.0) | 130 (83.9) | 33 (21.3) | 154 (99.4) |
| C1C2 | 174 (100) | 7.3 (1.5) | 159 (91.4) | 93 (53.4) | 160 (91.9) | 172 (98.8) | 122 (70.1) | 122 (70.1) | 93 (53.4) | 141 (81.0) | 30 (17.2) | 172 (98.8) |
| C2 | 61 (100) | 7.4 (1.4) | 55 (90.2) | 36 (59.0) | 54 (88.5) | 59 (96.7) | 49 (80.3)† | 50 (82.0) | 35 (57.4) | 47 (77.0) | 8 (13.1) | 60 (98.4) |
| **HLA-C -35 kb** | | | | | | | | | | | | |
| TT | 150 (100) | 7.2 (1.6) | 129 (86.0) | 82 (54.7) | 138 (92.0) | 145 (96.7) | 99 (66.0) | 105 (70.0) | 83 (55.3) | 127 (84.7) | 27 (18.0) | 149 (99.3) |
| CT | 202 (100) | 7.4 (1.3) | 181 (89.6) | 111 (54.9) | 182 (90.1) | 197 (97.5) | 146 (72.3) | 149 (73.8) | 125 (61.9) | 165 (81.7) | 32 (15.8) | 200 (99.0) |
| CC | 59 (100) | 7.4 (1.6) | 54 (91.5) | 33 (55.9) | 53 (89.8) | 58 (98.3) | 44 (74.6) | 50 (84.7)† | 29 (49.2) | 44 (75.6) | 13 (22.0) | 58 (98.3) |
| **NKG2Da** | | | | | | | | | | | | |
| CC | 270 (100) | 7.3 (1.4) | 238 (88.1) | 144 (53.3) | 246 (91.1) | 267 (98.9) | 183 (67.8) | 199 (73.7) | 150 (55.6) | 220 (81.5) | 51 (18.9) | 268 (99.2) |
| CT | 130 (100) | 7.4 (1.5) | 115 (88.5) | 76 (58.5) | 119 (91.5) | 123 (94.6)† | 97 (74.6) | 98 (75.4) | 75 (57.7) | 104 (80.0) | 22 (16.9) | 128 (98.5) |
| TT | 18 (100) | 7.4 (1.6) | 17 (94.4) | 10 (55.6) | 15 (83.3) | 17 (94.4) | 15 (83.3) | 13 (72.2) | 11 (61.1) | 16 (88.9) | 2 (11.1) | 18 (100) |
| **NKG2Db** | | | | | | | | | | | | |
| CC | 210 (100) | 7.4 (1.3) | 185 (88.1) | 121 (57.6) | 191 (90.9) | 208 (99.0)* | 147 (70.0) | 157 (74.8) | 119 (56.7) | 174 (82.8) | 39 (18.6) | 208 (99.0) |
| CG | 150 (100) | 7.3 (1.6) | 134 (89.3) | 78 (52.0) | 139 (92.7) | 143 (95.3)† | 107 (71.3) | 111 (74.0) | 86 (57.3) | 120 (80.0) | 29 (19.3) | 148 (98.7) |
| GG | 42 (100) | 7.2 (1.4) | 36 (85.7) | 24 (57.1) | 36 (85.7) | 40 (95.2) | 30 (71.4) | 30 (71.4) | 25 (59.5) | 36 (85.7) | 3 (7.1) | 42 (100) |

[a] Numbers may not sum to the overall total due to missing data.

[b] Percentages indicate the proportion of controls who are PyV seropositive versus PyV seronegative in each given strata.

[c] PyV infection was determined using seropositivity for the VP1 protein.

[d] All participants were seropositive for at least one of the ten PyVs assayed.

[e] Standard deviation referred to as "SD".

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ to test difference in proportions between groups, as determined by $X^2$ or Fisher's exact tests for categorical variables, and by Kruskal-Wallis or Wilcoxon rank sum tests for continuous variables.

† $P < 0.05$, †† $P < 0.01$, ††† $P < 0.005$ to test difference in proportions between groups, as determined by logistic regression predicting PyV serostatus by SNP genotype (reference group the low expression genotype), following adjustment for age group and gender.

Supplemental Table 3.6. Odds ratios and *P*-values for cutaneous squamous cell carcinoma (SCC) by quartiles of polyomavirus (PyV) seroreactivity, stratified by *HLA-C* genotype, among 713 study participants from the New Hampshire Skin Cancer Study.

| | C1 | P | C1C2 | P | C2 | P |
|---|---|---|---|---|---|---|
| BK Q2 | 1.28 | 0.50 | 0.71 | 0.36 | 3.13 | 0.16 |
| BK Q3 | 2.14 | **0.05** | 0.56 | 0.13 | 1.20 | 0.79 |
| BK Q4 | 1.68 | 0.19 | 0.88 | 0.73 | 1.57 | 0.55 |
| JC Q2 | 0.60 | 0.23 | 0.91 | 0.83 | 0.63 | 0.56 |
| JC Q3 | 0.92 | 0.83 | 1.32 | 0.46 | 1.00 | 1.00 |
| JC Q4 | 0.76 | 0.46 | 2.55 | **0.01** | 0.94 | 0.93 |
| KI Q2 | 1.49 | 0.30 | 1.44 | 0.34 | 1.53 | 0.54 |
| KI Q3 | 0.97 | 0.93 | 1.62 | 0.20 | 1.58 | 0.55 |
| KI Q4 | 0.93 | 0.84 | 0.97 | 0.94 | 2.01 | 0.35 |
| WU Q2 | 0.66 | 0.25 | 0.98 | 0.96 | 0.34 | 0.18 |
| WU Q3 | 0.63 | 0.22 | 1.30 | 0.45 | 0.65 | 0.52 |
| WU Q4 | 0.71 | 0.38 | 1.18 | 0.64 | 0.81 | 0.79 |
| MCV Q2 | 1.49 | 0.27 | 0.86 | 0.68 | 0.64 | 0.54 |
| MCV Q3 | 0.97 | 0.94 | 0.39 | **0.02** | 0.28 | 0.13 |
| MCV Q4 | 1.62 | 0.23 | 0.79 | 0.50 | 0.67 | 0.56 |
| HPyV6 Q2 | 1.24 | 0.57 | 1.74 | 0.16 | 0.84 | 0.80 |
| HPyV6 Q3 | 1.50 | 0.28 | 1.43 | 0.35 | 0.47 | 0.34 |
| HPyV6 Q4 | 1.24 | 0.57 | 1.32 | 0.44 | 1.10 | 0.90 |
| HPyV7 Q2 | 2.87 | **0.01** | 0.94 | 0.88 | 0.43 | 0.30 |
| HPyV7 Q3 | 1.30 | 0.54 | 0.90 | 0.79 | 0.26 | 0.06 |
| HPyV7 Q4 | 1.95 | 0.09 | 1.64 | 0.19 | 1.48 | 0.60 |
| TSV Q2 | 0.97 | 0.93 | 0.66 | 0.27 | 2.40 | 0.21 |
| TSV Q3 | 0.99 | 0.97 | 1.04 | 0.91 | 1.42 | 0.65 |
| TSV Q4 | 0.79 | 0.56 | 0.81 | 0.58 | 2.30 | 0.26 |
| HPyV9 Q2 | 1.31 | 0.48 | 1.17 | 0.68 | 1.01 | 0.99 |
| HPyV9 Q3 | 0.68 | 0.33 | 0.94 | 0.88 | 0.35 | 0.20 |
| HPyV9 Q4 | 1.11 | 0.78 | 1.54 | 0.24 | 1.81 | 0.42 |
| HPyV10 Q2 | 0.69 | 0.31 | 0.93 | 0.84 | 5.16 | **0.02** |
| HPyV10 Q3 | 0.86 | 0.71 | 0.68 | 0.29 | 1.73 | 0.52 |
| HPyV10 Q4 | 0.77 | 0.47 | 1.29 | 0.50 | 3.57 | 0.14 |

Supplemental Table 3.7. Odds ratios and *P*-values for cutaneous squamous cell carcinoma (SCC) by quartiles of polyomavirus (PyV) seroreactivity, stratified by *HLA-C -35 kb* genotype, among 713 study participants from the New Hampshire Skin Cancer Study.

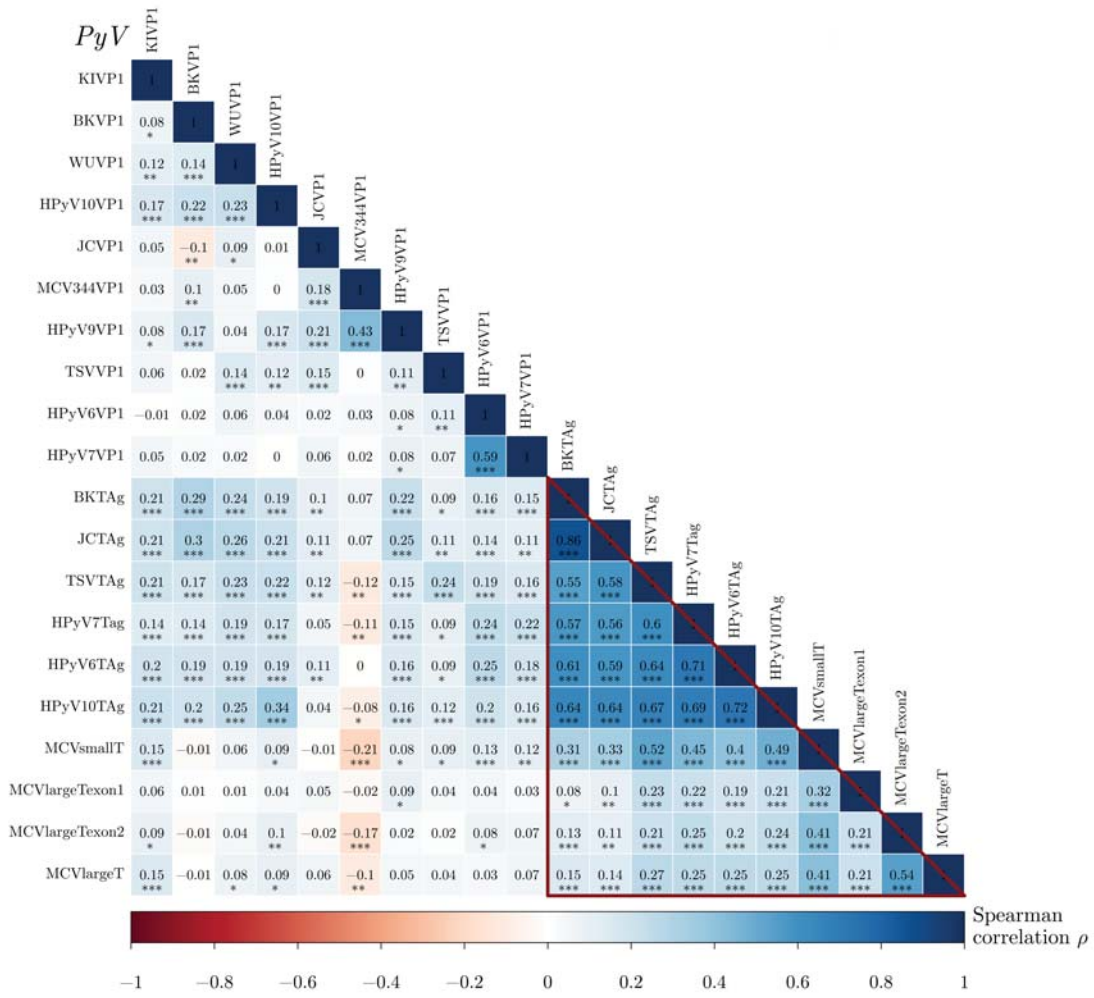
| | TT | P | CT | P | CC | P |
|---|---|---|---|---|---|---|
| BK Q2 | 0.86 | 0.71 | 0.97 | 0.93 | 1.66 | 0.40 |
| BK Q3 | 1.20 | 0.66 | 0.96 | 0.91 | 1.24 | 0.73 |
| BK Q4 | 2.00 | 0.08 | 0.89 | 0.74 | 0.83 | 0.81 |
| JC Q2 | 0.59 | 0.21 | 1.09 | 0.83 | 0.86 | 0.83 |
| JC Q3 | 0.87 | 0.73 | 1.36 | 0.38 | 2.37 | 0.21 |
| JC Q4 | 0.93 | 0.85 | 1.99 | 0.06 | 1.44 | 0.57 |
| KI Q2 | 1.54 | 0.26 | 1.59 | 0.19 | 2.25 | 0.21 |
| KI Q3 | 0.64 | 0.26 | 2.26 | **0.02** | 1.14 | 0.86 |
| KI Q4 | 0.51 | 0.11 | 1.52 | 0.23 | 1.97 | 0.40 |
| WU Q2 | 0.42 | **0.05** | 1.08 | 0.82 | 0.15 | **0.01** |
| WU Q3 | 0.84 | 0.64 | 1.28 | 0.46 | 0.47 | 0.25 |
| WU Q4 | 0.68 | 0.31 | 0.93 | 0.83 | 0.72 | 0.58 |
| MCV Q2 | 1.14 | 0.72 | 1.05 | 0.89 | 1.37 | 0.64 |
| MCV Q3 | 0.69 | 0.37 | 0.58 | 0.14 | 0.55 | 0.40 |
| MCV Q4 | 0.55 | 0.14 | 1.28 | 0.46 | 1.50 | 0.51 |
| HPyV6 Q2 | 1.73 | 0.17 | 1.05 | 0.89 | 0.54 | 0.39 |
| HPyV6 Q3 | 1.67 | 0.20 | 1.12 | 0.74 | 0.41 | 0.23 |
| HPyV6 Q4 | 1.31 | 0.53 | 1.20 | 0.58 | 0.51 | 0.35 |
| HPyV7 Q2 | 2.35 | **0.05** | 0.98 | 0.96 | 0.11 | **0.01** |
| HPyV7 Q3 | 1.13 | 0.79 | 0.65 | 0.22 | 0.40 | 0.20 |
| HPyV7 Q4 | 2.08 | 0.09 | 1.27 | 0.48 | 0.64 | 0.52 |
| TSV Q2 | 0.96 | 0.91 | 0.77 | 0.45 | 1.26 | 0.72 |
| TSV Q3 | 1.13 | 0.75 | 0.86 | 0.66 | 1.60 | 0.49 |
| TSV Q4 | 0.60 | 0.24 | 1.07 | 0.84 | 1.72 | 0.41 |
| HPyV9 Q2 | 0.75 | 0.44 | 2.35 | **0.02** | 1.20 | 0.81 |
| HPyV9 Q3 | 0.42 | **0.04** | 1.38 | 0.42 | 0.79 | 0.73 |
| HPyV9 Q4 | 0.52 | 0.10 | 3.30 | **0.00** | 1.42 | 0.55 |
| HPyV10 Q2 | 0.97 | 0.94 | 1.12 | 0.73 | 1.45 | 0.55 |
| HPyV10 Q3 | 0.99 | 0.99 | 0.58 | 0.13 | 1.99 | 0.29 |
| HPyV10 Q4 | 1.29 | 0.53 | 1.02 | 0.95 | 1.60 | 0.45 |

Supplemental Table 3.8. Odds ratios and *P*-values for cutaneous squamous cell carcinoma (SCC) by quartiles of polyomavirus (PyV) seroreactivity, stratified by *NKG2Da* genotype, among 713 study participants from the New Hampshire Skin Cancer Study.

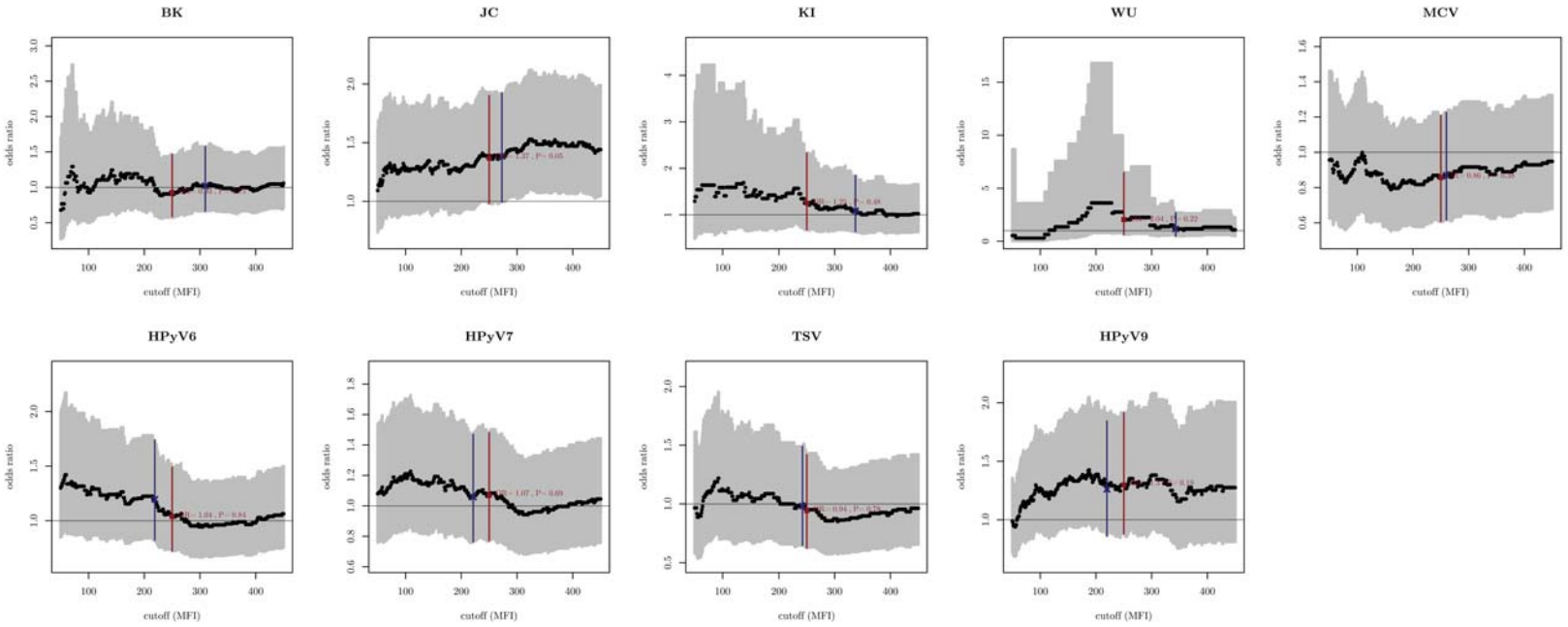
| | CC | P | CT | P | TT | P |
|---|---|---|---|---|---|---|
| BK Q2 | 1.02 | 0.96 | 0.94 | 0.88 | 0.00 | 1.00 |
| BK Q3 | 1.08 | 0.78 | 0.87 | 0.76 | 1.71 | 0.78 |
| BK Q4 | 1.01 | 0.98 | 1.01 | 0.99 | 2.83 | 0.52 |
| JC Q2 | 1.55 | 0.17 | 0.16 | **0.00** | 0.00 | 1.00 |
| JC Q3 | 1.84 | **0.04** | 0.58 | 0.23 | 86.41 | 0.10 |
| JC Q4 | 2.18 | **0.01** | 0.62 | 0.27 | 11.17 | 0.22 |
| KI Q2 | 1.83 | **0.04** | 1.61 | 0.29 | 1.00 | 1.00 |
| KI Q3 | 1.53 | 0.17 | 0.83 | 0.68 | 3980577597528562688.00 | 1.00 |
| KI Q4 | 1.08 | 0.81 | 0.85 | 0.72 | 2863504934.97 | 1.00 |
| WU Q2 | 0.79 | 0.43 | 0.45 | 0.09 | 0.00 | 1.00 |
| WU Q3 | 0.87 | 0.64 | 0.81 | 0.61 | 1.26 | 0.84 |
| WU Q4 | 1.15 | 0.62 | 0.55 | 0.16 | 0.51 | 0.67 |
| MCV Q2 | 1.22 | 0.49 | 1.00 | 0.99 | 0.70 | 0.83 |
| MCV Q3 | 0.80 | 0.47 | 0.31 | **0.04** | 0.10 | 0.22 |
| MCV Q4 | 1.26 | 0.41 | 0.72 | 0.47 | 0.20 | 0.30 |
| HPyV6 Q2 | 1.10 | 0.75 | 1.12 | 0.81 | 16.68 | 0.18 |
| HPyV6 Q3 | 1.31 | 0.36 | 0.91 | 0.84 | 1.68 | 0.77 |
| HPyV6 Q4 | 1.40 | 0.27 | 0.72 | 0.45 | 5.38 | 0.37 |
| HPyV7 Q2 | 1.03 | 0.92 | 0.92 | 0.85 | 0.20 | 0.50 |
| HPyV7 Q3 | 0.93 | 0.81 | 0.74 | 0.52 | 0.16 | 0.40 |
| HPyV7 Q4 | 1.70 | 0.07 | 0.89 | 0.80 | 0.93 | 0.97 |
| TSV Q2 | 0.83 | 0.53 | 1.29 | 0.58 | 7.51 | 0.41 |
| TSV Q3 | 1.05 | 0.87 | 1.16 | 0.75 | 2.68 | 0.56 |
| TSV Q4 | 1.09 | 0.79 | 1.04 | 0.93 | 0.55 | 0.75 |
| HPyV9 Q2 | 1.39 | 0.26 | 0.98 | 0.96 | 0.39 | 0.57 |
| HPyV9 Q3 | 0.77 | 0.40 | 0.91 | 0.85 | 0.26 | 0.48 |
| HPyV9 Q4 | 1.59 | 0.10 | 1.10 | 0.84 | 0.46 | 0.60 |
| HPyV10 Q2 | 1.17 | 0.59 | 1.02 | 0.97 | 3.62 | 0.44 |
| HPyV10 Q3 | 0.72 | 0.29 | 0.92 | 0.85 | 5.27 | 0.38 |
| HPyV10 Q4 | 0.97 | 0.92 | 1.33 | 0.55 | 1.34 | 0.88 |

Supplemental Table 3.9. Odds ratios and *P*-values for cutaneous squamous cell carcinoma (SCC) by quartiles of polyomavirus (PyV) seroreactivity, stratified by *NKG2Db* genotype, among 713 study participants from the New Hampshire Skin Cancer Study.

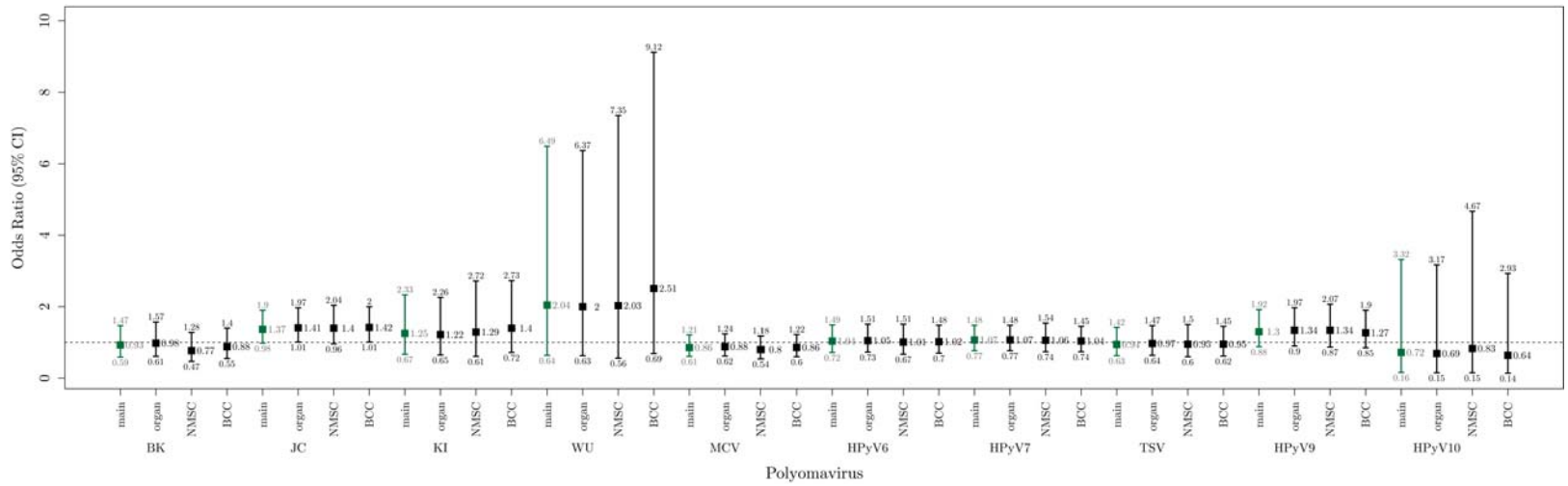
| | CC | P | CG | P | GG | P |
|---|---|---|---|---|---|---|
| BK Q2 | 1.08 | 0.81 | 1.07 | 0.86 | 0.77 | 0.80 |
| BK Q3 | 0.91 | 0.76 | 0.93 | 0.86 | 1.86 | 0.46 |
| BK Q4 | 0.85 | 0.62 | 1.65 | 0.23 | 1.91 | 0.47 |
| JC Q2 | 1.64 | 0.17 | 0.22 | **0.00** | 1.00 | 1.00 |
| JC Q3 | 1.47 | 0.26 | 1.15 | 0.71 | 4.86 | 0.16 |
| JC Q4 | 1.90 | 0.06 | 1.04 | 0.91 | 1.82 | 0.51 |
| KI Q2 | 1.58 | 0.16 | 1.36 | 0.44 | 4.12 | 0.11 |
| KI Q3 | 1.33 | 0.40 | 1.01 | 0.98 | 5.34 | 0.09 |
| KI Q4 | 1.20 | 0.59 | 0.80 | 0.60 | 2.94 | 0.26 |
| WU Q2 | 0.62 | 0.17 | 0.80 | 0.57 | 0.98 | 0.98 |
| WU Q3 | 0.97 | 0.93 | 0.56 | 0.13 | 2.64 | 0.22 |
| WU Q4 | 1.07 | 0.83 | 0.59 | 0.17 | 2.04 | 0.39 |
| MCV Q2 | 1.18 | 0.60 | 0.96 | 0.92 | 0.89 | 0.89 |
| MCV Q3 | 0.73 | 0.37 | 0.29 | **0.01** | 1.18 | 0.85 |
| MCV Q4 | 1.01 | 0.98 | 1.17 | 0.69 | 0.49 | 0.42 |
| HPyV6 Q2 | 1.12 | 0.74 | 1.16 | 0.73 | 1.53 | 0.64 |
| HPyV6 Q3 | 1.35 | 0.37 | 1.00 | 0.99 | 0.72 | 0.73 |
| HPyV6 Q4 | 1.17 | 0.64 | 1.12 | 0.78 | 1.08 | 0.93 |
| HPyV7 Q2 | 1.44 | 0.29 | 0.90 | 0.80 | 0.31 | 0.24 |
| HPyV7 Q3 | 1.02 | 0.96 | 0.81 | 0.62 | 0.20 | 0.10 |
| HPyV7 Q4 | 1.59 | 0.17 | 1.46 | 0.34 | 0.40 | 0.31 |
| TSV Q2 | 0.69 | 0.25 | 1.32 | 0.50 | 1.42 | 0.68 |
| TSV Q3 | 0.84 | 0.59 | 1.16 | 0.72 | 0.89 | 0.89 |
| TSV Q4 | 0.87 | 0.67 | 1.38 | 0.44 | 0.45 | 0.46 |
| HPyV9 Q2 | 1.68 | 0.12 | 1.22 | 0.62 | 0.46 | 0.39 |
| HPyV9 Q3 | 1.14 | 0.71 | 0.47 | 0.10 | 0.86 | 0.85 |
| HPyV9 Q4 | 1.75 | 0.09 | 1.24 | 0.58 | 0.67 | 0.67 |
| HPyV10 Q2 | 0.84 | 0.57 | 1.59 | 0.25 | 3.40 | 0.19 |
| HPyV10 Q3 | 0.58 | 0.12 | 1.09 | 0.83 | 4.07 | 0.20 |
| HPyV10 Q4 | 0.79 | 0.46 | 2.02 | 0.09 | 2.85 | 0.33 |



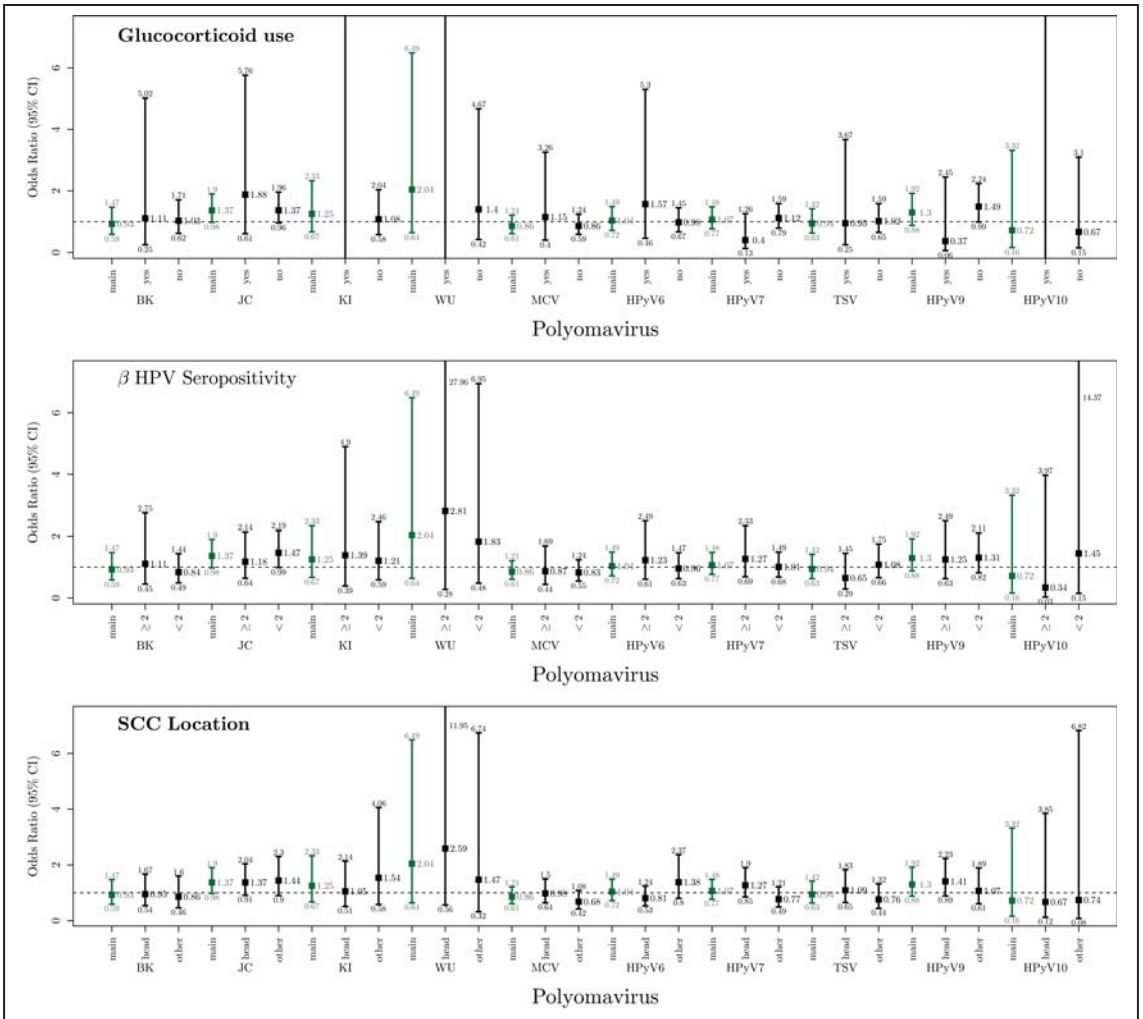

Supplemental Figure 3.1. Spearman rank correlation coefficients, ρ, between the median fluorescence intensity (MFI) values against each human polyomavirus (PyV) VP1 or T antigen (TAg) among 460 controls from the New Hampshire Skin Cancer Study, where * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Not all PyV TAgs were assayed, and MCV large TAg was assayed using the entire protein as well as with two fragments (exon 1 and exon 2) of the full length large TAg. The red triangle emphasizes the strong correlations between PyV TAgs.



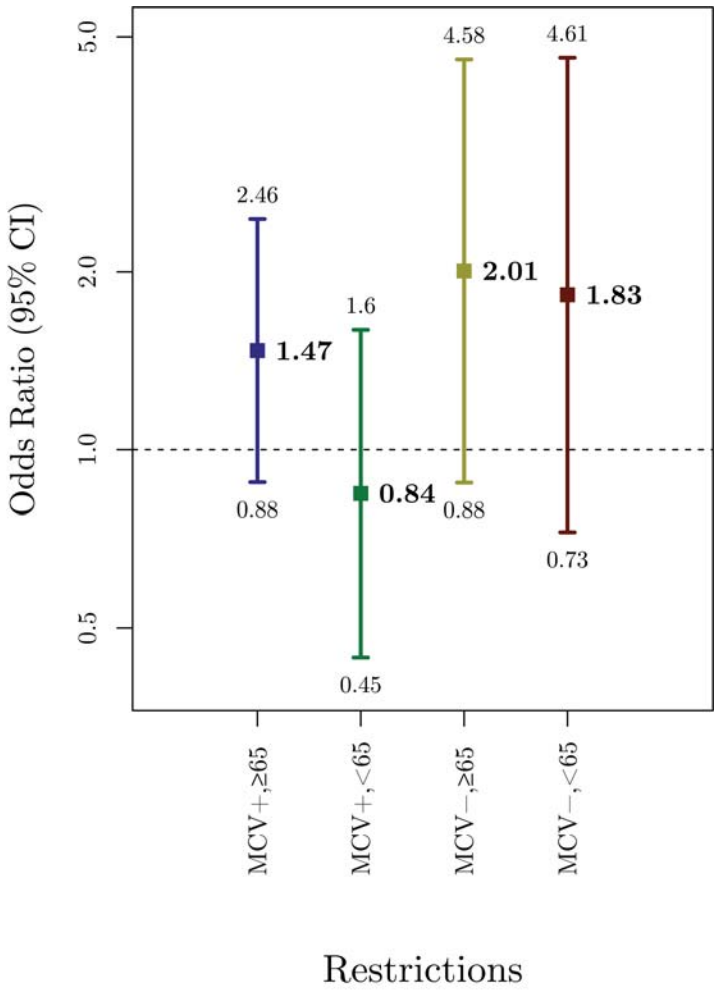

Supplemental Figure 3.2. Robustness of odds ratio (OR) estimates for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among 713 study participants from the New Hampshire Skin Cancer Study, following adjustment for age group and gender. The cut points were varied from 50 to 450 median fluorescence intensity (MFI) units (x axis), and the resulting ORs were calculated using the new cutoffs (y axis). The red dots show the ORs using the recommended cutoff of 250 MFI units. The blue dots denote the cut points calculated using a frequency distribution analysis described in van der Meijden *et al.*, 2013. OR estimates for HPyV10 could not be accurately computed due to the viruses' high seroprevalence. The gray bands are the 95% confidence intervals (CI) about each OR.




Supplemental Figure 3.3. Plot of odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among study participants from the New Hampshire Skin Cancer Study, when excluding participants with a history of organ transplantation ("organ", n=1 control and 6 cases), restricting to participants with no previous skin cancers ("NMSC", n=423 controls and 179 cases), and excluding participants with a concomitant basal cell carcinoma ("BCC", n=22 cases), following adjustment for age group and gender. "Main" refers to unstratified risk estimates presented in Figure 3.1. The dashed line represents an OR=1.




Supplemental Figure 3.4. Plot of odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among study participants from the New Hampshire Skin Cancer Study, when stratified by oral glucocorticoid use for one month or longer ("yes" refers to use (n=39 controls and 33 cases) and "no" to nonuse (n=415 controls and 211 cases); as people with a history of glucocorticoid use may also have undergone organ transplantation, we restricted the analysis to those who were not organ transplant recipients), β HPV seropositivity ("≥2" (n=125 controls and 82 cases) and "<2" (n=335 controls and 171 cases) refers to number of β HPV seropositive), and SCC location ("head" (n=146 cases) refers to SCC located on the head or neck, and "other" (n=97 cases) refers to SCC located on other body parts), following adjustment for age group and gender. "Main" refers to unstratified risk estimates presented in Figure 3.1. OR and 95% CI were not computed for strata in which all participants were seropositive for the PyV of interest (represented by a solid vertical black line). The dashed line represents an OR=1.




Supplemental Figure 3.5. Plot of odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for John Cunningham polyomavirus (JC) among study participants from the New Hampshire Skin Cancer Study, when restricting to participants with a combination of MCV serostatus (either seropositive or seronegative) and </≥ median age, following adjustment for age group and gender. The dashed line represents an OR=1.

Supplemental Figure 3.6. Plot of odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among study participants from the New Hampshire Skin Cancer Study, when stratified by *HLA-C* genotype (C1, C1C2, C2).

Supplemental Figure 3.7. Plot of odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among study participants from the New Hampshire Skin Cancer Study, when stratified by *HLA-C -35 kb* genotype (TT, CT, CC).

Supplemental Figure 3.8. Plot of odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among study participants from the New Hampshire Skin Cancer Study, when stratified by *NKG2Da* genotype (CC, CT, TT).

Supplemental Figure 3.9. Plot of odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among study participants from the New Hampshire Skin Cancer Study, when stratified by *NKG2Db* genotype (CC, CG, GG).

Supplemental Table 4.1. Distribution of BK and JC human polyomavirus (PyV) seropositivity by selected baseline characteristics among 229 controls from the Skin Cancer Prevention Study. [a]

| Characteristic | Total, No. (%) | Seropositive, No. (%) [b] | |
|---|---|---|---|
| | | BK | JC |
| Overall | 229 (100) | 222 (96.9) | 177 (77.3) |
| Gender | | | |
| Male | 201 (100) | 195 (97.0) | 154 (76.6) |
| Female | 28 (100) | 27 (96.4) | 23 (82.1) |
| Randomization arm in RCT | | | |
| Treatment | 115 (100) | 111 (96.5) | 88 (76.5) |
| Placebo | 114 (100) | 111 (97.4) | 89 (78.1) |
| Study center [c] | | | |
| DHMC | 43 (100) | 40 (93.0) | 34 (79.1) |
| UCLA | 65 (100) | 63 (96.9) | 50 (76.9) |
| UCSF | 56 (100) | 55 (98.2) | 42 (75.0) |
| UMN | 65 (100) | 64 (98.5) | 51 (78.5) |
| Previous skin cancers | | | |
| 1 | 102 (100) | 99 (97.1) | 79 (77.4) |
| 2 | 41 (100) | 40 (97.6) | 31 (75.6) |
| 3 | 20 (100) | 20 (100) | 17 (85.0) |
| 4-5 | 36 (100) | 35 (97.2) | 27 (75.0) |
| 6-9 | 17 (100) | 16 (94.1) | 15 (88.2) |
| ⩾10 | 11 (100) | 10 (90.9) | 7 (63.6) |
| Cigarette use | | | |
| Never smoked | 91 (100) | 89 (97.8) | 70 (76.9) |
| Former smoker | 111 (100) | 107 (96.4) | 86 (77.5) |
| Current smoker | 27 (100) | 26 (96.3) | 21 (77.8) |
| Body mass index (kg/m$^2$) | | | |
| Underweight <18.5 | 2 (100) | 2 (100) | 2 (100) |
| Normal 18.5-24.9 | 98 (100) | 93 (94.9) | 69 (70.4) |
| Overweight 45.0-29.9 | 108 (100) | 106 (98.1) | 89 (82.4) |
| Obese >30.0 | 15 (100) | 15 (100) | 13 (86.7) |
| Skin sun sensitivity | | | |
| Always or usually burns | 110 (100) | 106 (96.4) | 83 (75.4) |
| Burns moderately or minimally | 118 (100) | 115 (97.5) | 93 (78.8) |
| Extent of UV skin damage | | | |
| Mild | 62 (100) | 60 (96.8) | 48 (77.4) |
| Moderate | 134 (100) | 130 (97.0) | 103 (76.9) |
| Severe | 31 (100) | 30 (96.8) | 25 (80.6) |
| Sun bathed (hours) | | | |
| Never | 62 (100) | 58 (93.5) | 45 (72.6) |
| 0-200 | 63 (100) | 62 (98.4) | 53 (84.1) |
| 200-400 | 54 (100) | 54 (100) | 43 (79.6) |
| 400-600 | 33 (100) | 31 (93.9) | 21 (63.6) |
| >600 | 17 (100) | 17 (100) | 15 (88.2) |
| Occupational sun exposure (years) | | | |
| 0-7 | 78 (100) | 76 (97.4) | 62 (79.5) |
| 7-20 | 64 (100) | 62 (96.9) | 48 (75.0) |
| 21-40 | 40 (100) | 39 (97.5) | 26 (65.0) |
| >40 | 46 (100) | 44 (95.6) | 40 (86.9) |
| Eye color | | | |
| Blue, green, gray, hazel | 185 (100) | 178 (96.2) | 144 (77.8) |
| Brown, black | 44 (100) | 44 (100) | 33 (75.0) |
| Hair color | | | |
| Blonde, red | 61 (100) | 57 (93.4) | 43 (70.5) |
| Brown, black | 168 (100) | 165 (98.2) | 134 (79.8) |
| Vitamin use | | | |
| No | 128 (100) | 122 (95.3) | 97 (75.8) |
| Occasional | 37 (100) | 36 (97.3) | 28 (75.7) |
| Daily | 60 (100) | 60 (100) | 48 (80.0) |

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ to test difference in proportions between groups, as determined by $X^2$ or Fisher's exact tests for categorical variables, or by Kruskal-Wallis or Wilcoxon rank sum tests for continuous variables, and to test $P$ for trend across ordinal groups.

[a] Numbers may not sum to the overall total due to missing data.

[b] Percentages indicate the proportion of healthy adults who are PyV seropositive versus PyV seronegative in each given strata. PyV infection was determined in the baseline or earliest serum sample collected using seropositivity for the VP1 protein.

[c] This multicenter study was conducted at sites in California (University of California at Los Angeles School of Medicine (UCLA); University of California Medical School, San Francisco (UCSF)), Minnesota (University of Minnesota Schools of Medicine and Public Health, Minneapolis (UMN)), and New Hampshire (Dartmouth-Hitchcock Medical Center, Hanover (DHMC)), USA.

Supplemental Table 4.2. Unconditional odds ratios [a] (95% confidence intervals) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) [b] type and quartiles [c] of PyV seroreactivity at baseline among 342 study participants from the Skin Cancer Prevention Study.

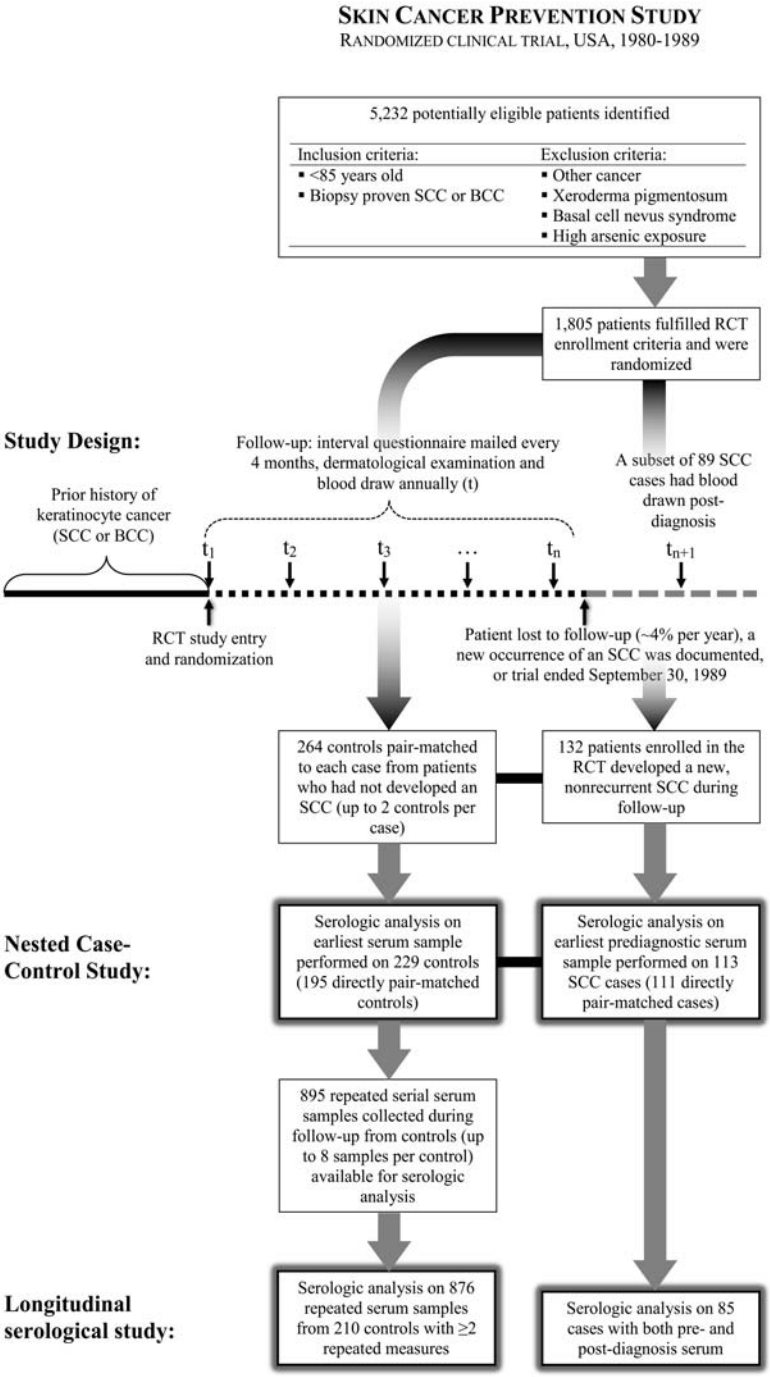
| PyV seroreactivity (MFI units) | Controls (n=229), No. (%) | SCC Cases (n=113) | |
|---|---|---|---|
| | | No. (%) | OR (95% CI) |
| BK | | | |
| Seronegative | 7 (3.1) | 1 (0.9) | 1.00 (referent) |
| Seropositive | 222 (96.9) | 112 (99.1) | 3.67 (0.44-30.42) |
| Quartile 1 | 57 (24.9) | 26 (23.0) | 1.00 (referent) |
| Quartile 2 | 57 (24.9) | 15 (13.3) | 0.60 (0.28-1.26) |
| Quartile 3 | 57 (24.9) | 33 (29.2) | 1.36 (0.71-2.61) |
| Quartile 4 | 58 (25.3) | 39 (34.5) | 1.57 (0.84-2.95) |
| *P* for trend [d] | | | 0.041 |
| JC | | | |
| Seronegative | 52 (22.7) | 13 (11.5) | 1.00 (referent) |
| Seropositive | 177 (77.3) | 100 (88.5) | 2.27 (1.18-4.40) |
| Quartile 1 | 57 (24.9) | 18 (15.9) | 1.00 (referent) |
| Quartile 2 | 58 (25.3) | 27 (23.9) | 1.51 (0.74-3.05) |
| Quartile 3 | 57 (24.9) | 22 (19.5) | 1.22 (0.59-2.52) |
| Quartile 4 | 57 (24.9) | 46 (40.7) | 2.62 (1.34-5.11) |
| *P* for trend [d] | | | 0.0078 |

[a] Adjusted for continuous age, gender, and study center. OR=odds ratios obtained from unconditional logistic regression analysis with adjustment for matching factors, CI=confidence interval.

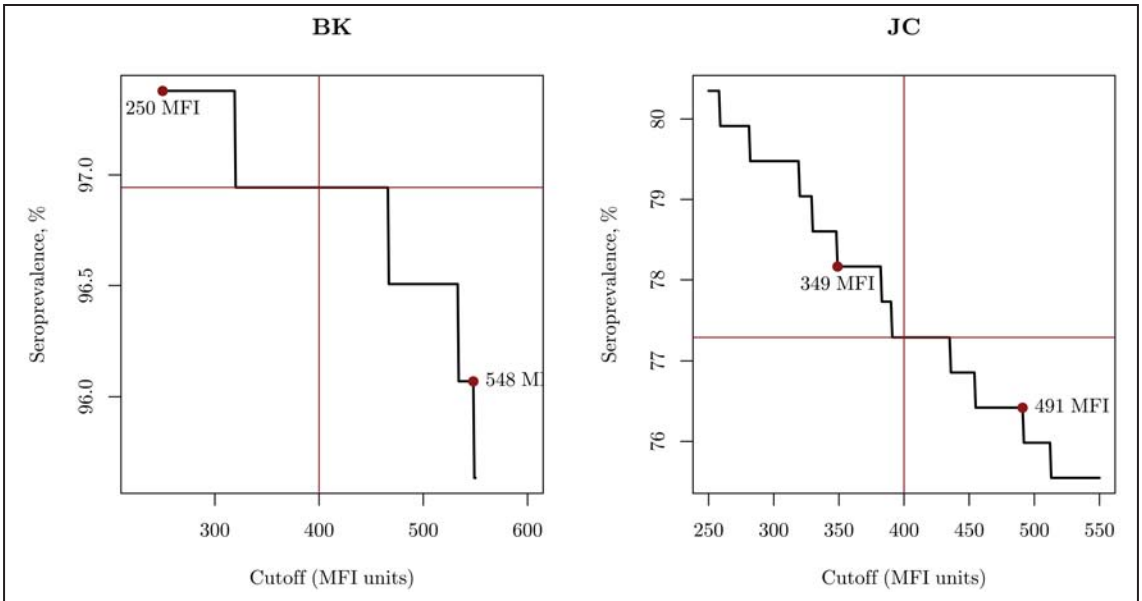
[b] PyV infection was determined in the baseline or earliest serum sample collected using seropositivity for the VP1 protein.

[c] Controls may not be evenly distributed within quartiles due to uneven data distribution.

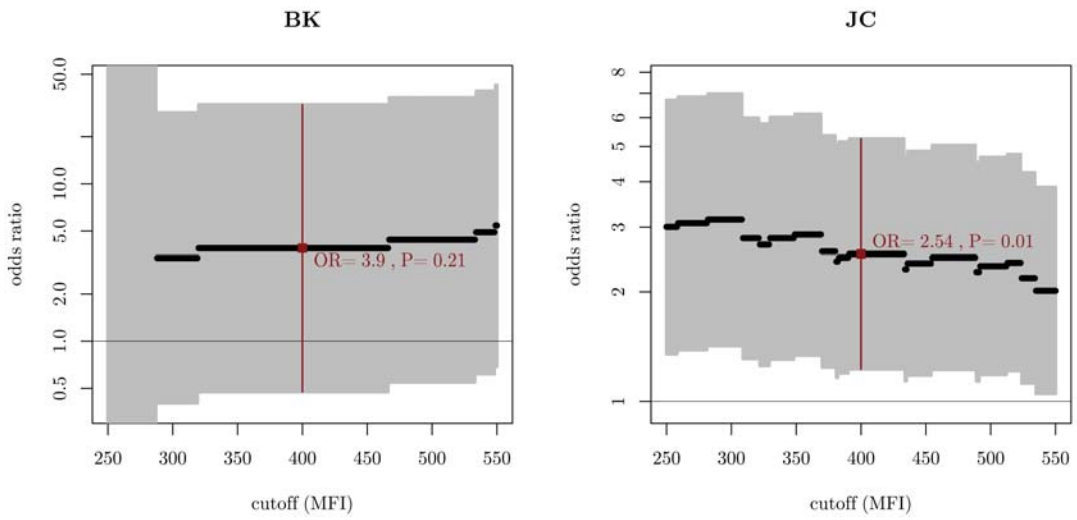
[d] Based on the seroreactivity quartiles modelled as a continuous variable.



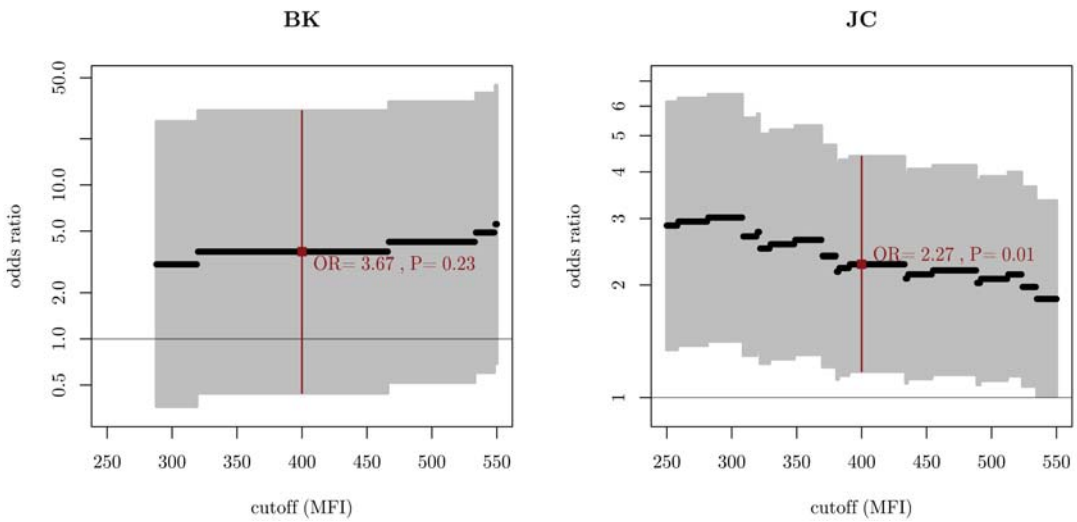

Supplemental Figure 4.1. Flowchart of the "Skin Cancer Prevention Study" randomized clinical trial (RCT) design from which the study group was derived. Patients with a prior history of squamous cell carcinoma (SCC) or basal cell carcinoma (BCC) were enrolled. In a nested case-control study, risk of a new, nonrecurrent SCC was associated with polyomavirus (PyV) seropositivity assessed in the earliest pre-diagnostic serum sample available (n=113 cases and 229 controls). Further, repeated serum measures drawn from controls (n=210 controls with 876 serum samples), and serum samples drawn pre- and post-SCC diagnosis from cases (n=85 cases), were investigated to determine the serostability of PyV antibodies over time.



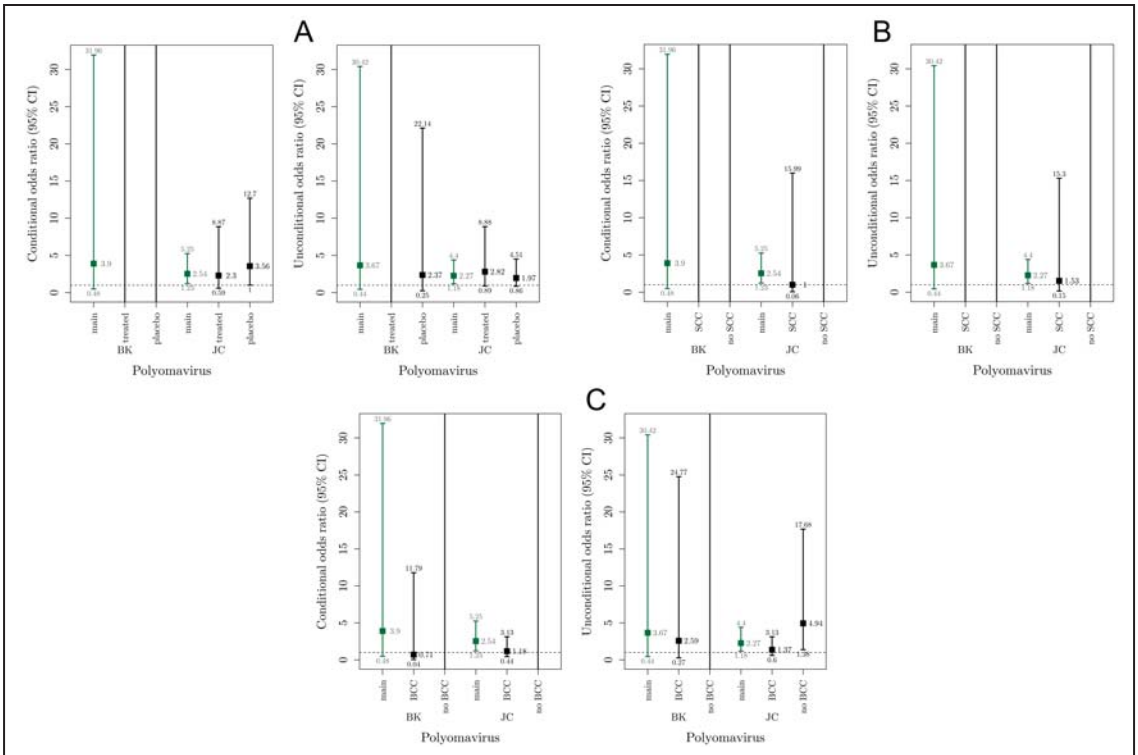

Supplemental Figure 4.2. Robustness of human polyomavirus (PyV) seroprevalence estimates among 229 controls from the Skin Cancer Prevention Study. The cut points were varied from 250 to 550 median fluorescence intensity (MFI) units (x axis), and the resulting seroprevalences were calculated using the new cutoffs (y axis). The red cross hairs show the seroprevalences for each PyV VP1 using the recommended cutoff=400 MFI units. The two red dots show the range of MFI units that would result in seroprevalences ±1% of the seroprevalences calculated using the recommended cutoff.



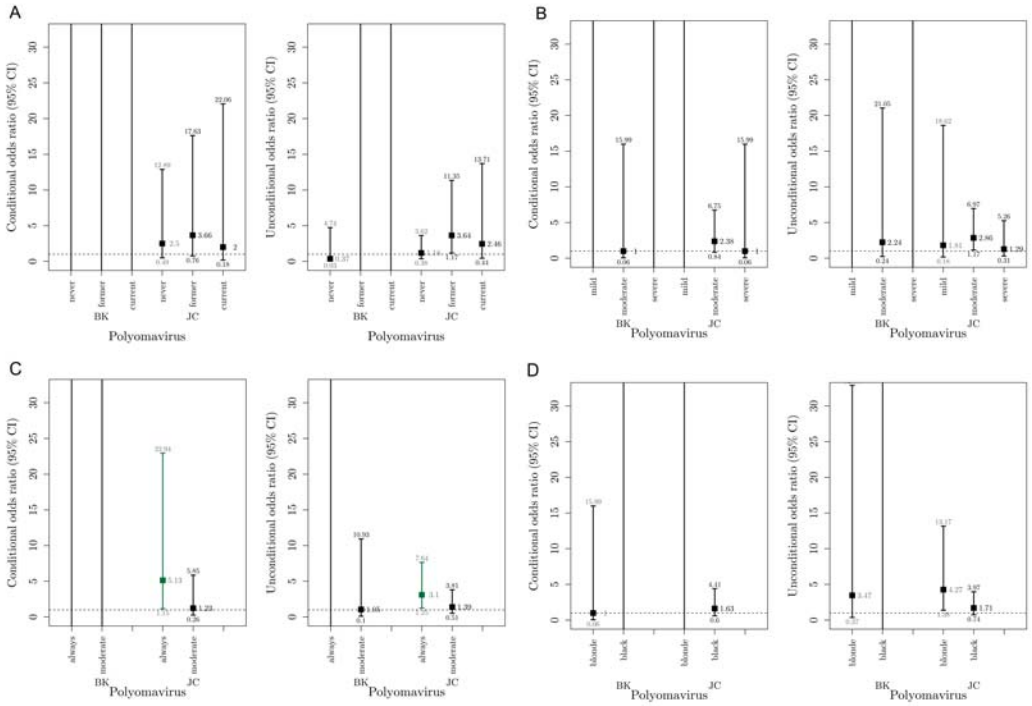

Supplemental Figure 4.3. Robustness of odds ratio (OR) estimates obtained from conditional logistic regression analysis for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among 306 study participants from the Skin Cancer Prevention Study. The cut points were varied from 250 to 550 median fluorescence intensity (MFI) units (x axis), and the resulting ORs were calculated using the new cutoffs (y axis). The red dots show the ORs using the recommended cutoff of 400 MFI units. OR estimates for BK could not be accurately computed below ~275 MFI units due to the viruses' high seroprevalence. The gray bands are the 95% confidence intervals (CI) about each OR.




Supplemental Figure 4.4. Robustness of odds ratio (OR) estimates obtained from unconditional logistic regression analysis for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among 342 study participants from the Skin Cancer Prevention Study, with adjustment for continuous age, gender, and study center. The cut points were varied from 250 to 550 median fluorescence intensity (MFI) units (x axis), and the resulting ORs were calculated using the new cutoffs (y axis). The red dots show the ORs using the recommended cutoff of 400 MFI units. OR estimates for BK could not be accurately computed below ~275 MFI units due to the viruses' high seroprevalence. The gray bands are the 95% confidence intervals (CI) about each OR.




Supplemental Figure 4.5. Plot of conditional (left) and unconditional (right) odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus type among study participants from the Skin Cancer Prevention Study, when stratified by **A.** treatment arm of the randomized clinical trial from which the participants were drawn ("treated" refers to patient assignment to the β-carotene treatment group (n=99 controls and 64 cases for conditional analyses; n=115 controls and 66 cases for unconditional analyses) and "placebo" to the placebo group (n=96 controls and 47 cases for conditional analyses; n=114 controls and 47 cases for unconditional analyses)), **B.** having ever had a prior SCC ("SCC" refers to having had a prior SCC (n=9 controls and 6 cases for conditional analyses; n=9 controls and 6 cases for unconditional analyses) and "no SCC" refers to never having had a prior SCC (n=186 controls and 105 cases for conditional analyses; n=220 controls and 107 cases for unconditional analyses)), and **C.** having ever had a prior BCC ("BCC" refers to having had a prior BCC (n=100 controls and 68 cases for conditional analyses; n=118 controls and 69 cases for unconditional analyses) and "no BCC" refers to never having had a prior BCC (n=95 controls and 43 cases for conditional analyses; n=111 controls and 44 cases for unconditional analyses)). Unconditional odds ratios were adjusted for continuous age, gender, and study center. "Main" refers to unstratified risk estimates presented in Table 4.4 and Supplemental Table 4.2. OR and 95% CI were not computed for uninformative pairs or strata in which all participants were seropositive for the PyV of interest (represented by a solid vertical black line). The dashed line represents an OR=1.




Supplemental Figure 4.6. Plot of conditional (left) and unconditional (right) odds ratios (95% confidence intervals as whiskers) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus type among study participants from the Skin Cancer Prevention Study, when stratified (as delineated in Table 4.1) by **A.** smoking status ("never" refers to never smokers, "former" refers to former smokers, and "current" refers to current smokers), **B.** UV skin damage ("mild" refers to mild skin damage, "moderate" refers to moderate skin damage, and "severe" refers to severe skin damage), **C.** skin sun sensitivity ("always" refers to always/usually burns and "moderate" refers to burns moderately/minimally), and **D.** hair color ("blonde" refers to blonde/red hair and "black" refers to brown/black hair). Unconditional odds ratios were adjusted for continuous age, gender, and study center. OR and 95% CI were not computed for uninformative pairs or strata in which all participants were seropositive for the PyV of interest (represented by a solid vertical black line). The dashed line represents an OR=1.

# References


1. Khalili K, Stoner GL. *Human polyomaviruses: molecular and clinical perspectives*. Wiley-Liss, Inc., 2001.
2. Randhawa P, Vats A, Shapiro R. The pathobiology of polyomavirus infection in man. *Polyomaviruses and Human Diseases*. Springer, 2006, pp 148-159.
3. Moens U, Ludvigsen M, Van Ghelue M. Human polyomaviruses in skin diseases. *Pathology research international* 2011, **2011**.
4. DeCaprio JA, Garcea RL. A cornucopia of human polyomaviruses. *Nature Reviews Microbiology* 2013, **11**(4)**:** 264-276.
5. Kean JM, Rao S, Wang M, Garcea RL. Seroepidemiology of human polyomaviruses. *PLoS Pathogens* 2009, **5**(3)**:** e1000363.
6. Ehlers B, Wieland U. The novel human polyomaviruses HPyV6, 7, 9 and beyond. *Apmis* 2013, **121**(8)**:** 783-795.
7. Jiang M, Abend JR, Johnson SF, Imperiale MJ. The role of polyomaviruses in human disease. *Virology* 2009, **384**(2)**:** 266-273.
8. Imperiale MJ. The human polyomaviruses, BKV and JCV: molecular pathogenesis of acute disease and potential role in cancer. *Virology* 2000, **267**(1)**:** 1-7.
9. van der Meijden E, Janssens RW, Lauber C, Bavinck JNB, Gorbalenya AE, Feltkamp MC. Discovery of a new human polyomavirus associated with trichodysplasia spinulosa in an immunocompromized patient. *PLoS Pathogens* 2010, **6**(7)**:** e1001024.
10. Schwieger-Briel A, Balma-Mena A, Ngan B, Dipchand A, Pope E. Trichodysplasia spinulosa—a rare complication in immunosuppressed patients. *Pediatric Dermatology* 2010, **27**(5)**:** 509-513.
11. Haycox CL, Kim S, Fleckman P, Smith LT, Piepkorn M, Sundberg JP*, et al.* Trichodysplasia spinulosa–a newly described folliculocentric viral infection in an immunocompromised host. Journal of Investigative Dermatology Symposium Proceedings; 1999: Nature Publishing Group; 1999. p. 268-271.
12. Shuda M, Feng H, Kwun HJ, Rosen ST, Gjoerup O, Moore PS*, et al.* T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. *Proceedings of the National Academy of Sciences* 2008, **105**(42)**:** 16272-16277.
13. Reisinger DM, Shiffer JD, Cognetta AB, Jr., Chang Y, Moore PS. Lack of evidence for basal or squamous cell carcinoma infection with Merkel cell polyomavirus in immunocompetent patients with Merkel cell carcinoma. *Journal of the American Academy of Dermatology* 2010, **63**(3)**:** 400-403.
14. Carter JJ, Paulson KG, Wipf GC, Miranda D, Madeleine MM, Johnson LG*, et al.* Association of Merkel cell polyomavirus–specific antibodies with Merkel cell carcinoma. *Journal of the National Cancer Institute* 2009, **101**(21)**:** 1510-1522.
15. De Luca A, Baldi A, Esposito V, Howard CM, Bagella L, Rizzo P*, et al.* The retinoblastoma gene family pRb/p105, p107, pRb2/p130 and simian virus-40 large T-antigen in human mesotheliomas. *Nature Medicine* 1997, **3**(8)**:** 913-916.
16. Carbone M, Rizzo P, Grimley PM, Procopio A, Mew DJ, Shridhar V*, et al.* Simian virus-40 large-T antigen binds p53 in human mesotheliomas. *Nature Medicine* 1997, **3**(8)**:** 908-912.

17. Saribas AS, Coric P, Hamazaspyan A, Davis W, Axman R, White MK*, et al.* Emerging from the Unknown: Structural and Functional Features of Agnoprotein of Polyomaviruses. *Journal of Cellular Physiology* 2016.
18. Gross L. A filterable agent, recovered from Ak leukemic extracts, causing salivary gland carcinomas in C3H mice. *Experimental Biology and Medicine* 1953, **83**(2)**:** 414-421.
19. Stoner GL, Hübner R. The human polyomaviruses: past, present, and future. *Human polyomaviruses: molecular and clinical perspectives* 2001**:** 611-663.
20. Padgett B, Zurhein G, Walker D, Eckroade R, Dessel B. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *The Lancet* 1971, **297**(7712)**:** 1257-1260.
21. Gardner S, Field A, Coleman D, Hulme B. New human papovavirus (BK) isolated from urine after renal transplantation. *The Lancet* 1971, **297**(7712)**:** 1253-1257.
22. Allander T, Andreasson K, Gupta S, Bjerkner A, Bogdanovic G, Persson MA*, et al.* Identification of a third human polyomavirus. *Journal of virology* 2007, **81**(8)**:** 4130-4136.
23. Gaynor AM, Nissen MD, Whiley DM, Mackay IM, Lambert SB, Wu G*, et al.* Identification of a novel polyomavirus from patients with acute respiratory tract infections. *PLoS pathogens* 2007, **3**(5)**:** e64.
24. Feng H, Shuda M, Chang Y, Moore PS. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 2008, **319**(5866)**:** 1096-1100.
25. Rodig SJ, Cheng J, Wardzala J, DoRosario A, Scanlon JJ, Laga AC*, et al.* Improved detection suggests all Merkel cell carcinomas harbor Merkel polyomavirus. *The Journal of clinical investigation* 2012, **122**(12)**:** 4645.
26. Schmitt M, Wieland U, Kreuter A, Pawlita M. C-terminal deletions of Merkel cell polyomavirus large T-antigen, a highly specific surrogate marker for virally induced malignancy. *International Journal of Cancer* 2012, **131**(12)**:** 2863-2868.
27. Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell host & microbe* 2010, **7**(6)**:** 509-515.
28. Scuda N, Hofmann J, Calvignac-Spencer S, Ruprecht K, Liman P, Kühn J*, et al.* A novel human polyomavirus closely related to the African green monkey-derived lymphotropic polyomavirus. *Journal of virology* 2011, **85**(9)**:** 4586-4590.
29. Sauvage V, Foulongne V, Cheval J, Gouilh MA, Pariente K, Dereure O*, et al.* Human polyomavirus related to African green monkey lymphotropic polyomavirus. *Emerging infectious diseases* 2011, **17**(8).
30. Buck CB, Phan GQ, Raiji MT, Murphy PM, McDermott DH, McBride AA. Complete genome sequence of a tenth human polyomavirus. *Journal of virology* 2012, **86**(19)**:** 10887-10887.
31. Feltkamp MC, Kazem S, van der Meijden E, Lauber C, Gorbalenya AE. From Stockholm to Malawi: recent developments in studying human polyomaviruses. *Journal of General Virology* 2013, **94**(Pt 3)**:** 482-496.
32. Siebrasse EA, Reyes A, Lim ES, Zhao G, Mkakosya RS, Manary MJ*, et al.* Identification of MW polyomavirus, a novel polyomavirus in human stool. *Journal of virology* 2012, **86**(19)**:** 10321-10326.

33. Yu G, Greninger AL, Isa P, Phan TG, Martínez MA, de la Luz Sanchez M, *et al.* Discovery of a novel polyomavirus in acute diarrheal samples from children. *PloS One* 2012, **7**(11): e49449.
34. Stolt A, Sasnauskas K, Koskela P, Lehtinen M, Dillner J. Seroepidemiology of the human polyomaviruses. *Journal of General Virology* 2003, **84**(6): 1499-1504.
35. Egli A, Infanti L, Dumoulin A, Buser A, Samaridis J, Stebler C, *et al.* Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors. *Journal of Infectious Diseases* 2009, **199**(6): 837-846.
36. Gardner SD. Prevalence in England of antibody to human polyomavirus (BK). *British medical journal* 1973, **1**(5845): 77.
37. Šroller V, Hamšíková E, Ludvíková V, Vochozková P, Kojzarová M, Fraiberk M, *et al.* Seroprevalence rates of BKV, JCV, and MCPyV polyomaviruses in the general Czech Republic population. *Journal of medical virology* 2013.
38. Viscidi RP, Rollison DE, Sondak VK, Silver B, Messina JL, Giuliano AR, *et al.* Age-specific seroprevalence of Merkel cell polyomavirus, BK virus, and JC virus. *Clinical and Vaccine Immunology* 2011, **18**(10): 1737-1743.
39. Antonsson A, Green AC, Mallitt K-A, O'Rourke PK, Pawlita M, Waterboer T, *et al.* Prevalence and stability of antibodies to the BK and JC polyomaviruses: a long-term longitudinal study of Australians. *Journal of General Virology* 2010, **91**(7): 1849-1853.
40. Nguyen NL, Le B-M, Wang D. Serologic evidence of frequent human infection with WU and KI polyomaviruses. *Emerging infectious diseases* 2009, **15**(8).
41. Chen T, Mattila PS, Jartti T, Ruuskanen O, Söderlund-Venermo M, Hedman K. Seroepidemiology of the newly found trichodysplasia spinulosa–associated polyomavirus. *Journal of Infectious Diseases* 2011, **204**(10): 1523-1526.
42. Neske F, Prifert C, Scheiner B, Ewald M, Schubert J, Opitz A, *et al.* High prevalence of antibodies against polyomavirus WU, polyomavirus KI, and human bocavirus in German blood donors. *BMC infectious diseases* 2010, **10**(1): 215.
43. Moens U, Van Ghelue M, Song X, Ehlers B. Serological cross-reactivity between human polyomaviruses. *Reviews in medical virology* 2013, **23**(4): 250-264.
44. Moens U, Johannessen M, Bárcena-Panero A, Gerits N, Ghelue Mv. Emerging polyomaviruses in the human population. *Reviews in Infection* 2010, **1**(2): 59-93.
45. Zhang N-N, Zhao L-Q, Qian Y, Zhu R-N, Deng J, Wang F, *et al.* Common WU polyomavirus infection in a Beijing population indicated by surveillance for serum IgG antibody against capsid protein VP2. *World Journal of Pediatrics* 2013, **9**(1): 48-52.
46. Nicol JT, Touzé A, Robinot R, Arnold F, Mazzoni E, Tognon M, *et al.* Seroprevalence and cross-reactivity of human polyomavirus 9. *Emerging infectious diseases* 2012, **18**(8): 1329.
47. Pastrana DV, Tolstov YL, Becker JC, Moore PS, Chang Y, Buck CB. Quantitation of human seroresponsiveness to Merkel cell polyomavirus. *PLoS pathogens* 2009, **5**(9): e1000578.
48. Nicol JT, Robinot R, Carpentier A, Carandina G, Mazzoni E, Tognon M, *et al.* Age-specific seroprevalences of merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and trichodysplasia spinulosa-associated polyomavirus. *Clinical and Vaccine Immunology* 2013, **20**(3): 363-368.

49. Pastrana DV, Wieland U, Silling S, Buck CB, Pfister H. Positive correlation between Merkel cell polyomavirus viral load and capsid-specific antibody titer. *Medical microbiology and immunology* 2012, **201**(1): 17-23.
50. van der Meijden E, Kazem S, Burgers MM, Janssens R, Bavinck JNB, de Melker H, *et al.* Seroprevalence of trichodysplasia spinulosa–associated polyomavirus. *Emerging infectious diseases* 2011, **17**(8): 1355.
51. Berrios C, Jung J, Primi B, Wang M, Pedamallu C, Duke F, *et al.* Malawi polyomavirus is a prevalent human virus that interacts with known tumor suppressors. *Journal of virology* 2014.
52. Nicol JT, Leblond V, Arnold F, Guerra G, Mazzoni E, Tognon M, *et al.* Seroprevalence of human Malawi Polyomavirus. *Journal of clinical microbiology* 2014, **52**(1): 321-323.
53. Kjærheim K, Røe OD, Waterboer T, Sehr P, Rizk R, Dai HY, *et al.* Absence of SV40 antibodies or DNA fragments in prediagnostic mesothelioma serum samples. *International journal of cancer* 2007, **120**(11): 2459-2465.
54. Rollison DE, Giuliano AR, Messina JL, Fenske NA, Cherpelis BS, Sondak VK, *et al.* Case–control study of Merkel cell polyomavirus infection and cutaneous squamous cell carcinoma. *Cancer Epidemiology Biomarkers & Prevention* 2012, **21**(1): 74-81.
55. Sehr P, Müller M, Höpfl R, Widschwendter A, Pawlita M. HPV antibody detection by ELISA with capsid protein L1 fused to glutathione S-transferase. *Journal of virological methods* 2002, **106**(1): 61-70.
56. Michael KM, Waterboer T, Sehr P, Rother A, Reidel U, Boeing H, *et al.* Seroprevalence of 34 human papillomavirus types in the German general population. *PLoS pathogens* 2008, **4**(6): e1000091.
57. Sehr P, Zumbach K, Pawlita M. A generic capture ELISA for recombinant proteins fused to glutathione S-transferase: validation for HPV serology. *Journal of immunological methods* 2001, **253**(1): 153-162.
58. Waterboer T, Sehr P, Michael KM, Franceschi S, Nieland JD, Joos TO, *et al.* Multiplex human papillomavirus serology based on in situ–purified glutathione S-transferase fusion proteins. *Clinical chemistry* 2005, **51**(10): 1845-1853.
59. Waterboer T, Sehr P, Pawlita M. Suppression of non-specific binding in serological Luminex assays. *Journal of immunological methods* 2006, **309**(1): 200-204.
60. Waterboer T, Neale R, Michael KM, Sehr P, de Koning MN, Weißenborn SJ, *et al.* Antibody responses to 26 skin human papillomavirus types in the Netherlands, Italy and Australia. *Journal of General Virology* 2009, **90**(8): 1986-1998.
61. Farzan SF, Waterboer T, Gui J, Nelson HH, Li Z, Michael KM, *et al.* Cutaneous alpha, beta and gamma human papillomaviruses in relation to squamous cell carcinoma of the skin: A population-based study. *International Journal of Cancer* 2013, **133**(7): 1713-1720.
62. Karagas MR, Nelson HH, Sehr P, Waterboer T, Stukel TA, Andrew A, *et al.* Human papillomavirus infection and incidence of squamous cell and basal cell carcinomas of the skin. *Journal of the National Cancer Institute* 2006, **98**(6): 389-395.
63. Topalis D, Andrei G, Snoeck R. The large tumor antigen: A "Swiss Army knife" protein possessing the functions required for the polyomavirus life cycle. *Antiviral research* 2013, **97**(2): 122-136.
64. Pipas JM. SV40: Cell transformation and tumorigenesis. *Virology* 2009, **384**(2): 294-303.

65. Rundell K, Major E, Lampert M. Association of cellular 56,000-and 32,000-molecular-weight protein with BK virus and polyoma virus T-antigens. *Journal of virology* 1981, **37**(3): 1090.
66. Sontag E, Fedorov S, Kamibayashi C, Robbins D, Cobb M, Mumby M. The interaction of SV40 small tumor antigen with protein phosphatase 2A stimulates the map kinase pathway and induces cell proliferation. *Cell* 1993, **75**(5): 887-897.
67. Darbinyan A, Darbinian N, Safak M, Radhakrishnan S, Giordano A, Khalili K. Evidence for dysregulation of cell cycle by human polyomavirus, JCV, late auxiliary protein. *Oncogene* 2002, **21**(36): 5574-5581.
68. Darbinyan A, White MK, Akan S, Radhakrishnan S, Del Valle L, Amini S, *et al.* Alterations of DNA damage repair pathways resulting from JCV infection. *Virology* 2007, **364**(1): 73-86.
69. Del Valle L, Khalili K. Detection of human polyomavirus proteins, T-antigen and agnoprotein, in human tumor tissue arrays. *Journal of medical virology* 2010, **82**(5): 806-811.
70. Moens U. Silencing viral microRNA as a novel antiviral therapy? *BioMed Research International* 2009, **2009**.
71. Hampras SS, Michel A, Schmitt M, Waterboer T, Kranz L, Gheit T, *et al.* Merkel cell polyomavirus (MCV) T-antigen seroreactivity, MCV DNA in eyebrow hairs, and squamous cell carcinoma. *Infectious agents and cancer* 2015, **10**(1): 1-5.
72. Chesters PM, Heritage J, McCance DJ. Persistence of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues. *Journal of Infectious Diseases* 1983, **147**(4): 676-684.
73. Heritage J, Chesters PM, McCance DJ. The persistence of papovavirus BK DNA sequences in normal human renal tissue. *Journal of medical virology* 1981, **8**(2): 143-150.
74. Monini P, Rotola A, Di Luca D, De Lellis L, Chiari E, Corallini A, *et al.* DNA rearrangements impairing BK virus productive infection in urinary tract tumors. *Virology* 1995, **214**(1): 273-279.
75. Shah KV, Daniel RW, Stone KR, Elliott AY. Investigation of human urogenital tract tumors for papovavirus etiology: brief communication. *Journal of the National Cancer Institute* 1978, **60**(3): 579-582.
76. Geetha D, Tong BC, Racusen L, Markowitz JS, Westra WH. Bladder carcinoma in a transplant recipient: evidence to implicate the BK human polyomavirus as a causal transforming agent. *Transplantation* 2002, **73**(12): 1933-1936.
77. Dalianis T, Hirsch HH. Human polyomaviruses in disease and cancer. *Virology* 2013, **437**(2): 63-72.
78. Abend JR, Jiang M, Imperiale MJ. BK virus and human cancer: innocent until proven guilty. *Seminars in cancer biology* 2009, **19**(4): 252-260.
79. Pino L, Rijo E, Nohales G, Frances A, Ubre A, Arango O. Bladder transitional cell carcinoma and BK virus in a young kidney transplant recipient. *Transplant Infectious Disease* 2013, **15**(1): E25-E27.
80. Vilkin A, Ronen Z, Levi Z, Morgenstern S, Halpern M, Niv Y. Presence of JC virus DNA in the tumor tissue and normal mucosa of patients with sporadic colorectal cancer (CRC) or with positive family history and Bethesda criteria. *Digestive diseases and sciences* 2012, **57**(1): 79-84.

81. Monini P, Rotola A, de Lellis L, Corallini A, Secchiero P, Albini A*, et al.* Latent BK virus infection and Kaposi's sarcoma pathogenesis. *International journal of Câncer* 1996, **66**(6)**:** 717-722.
82. Giraud G, Ramqvist T, Ragnarsson-Olding B, Dalianis T. DNA from BK virus and JC virus and from KI, WU, and MC polyomaviruses as well as from simian virus 40 is not detected in non-UV-light-associated primary malignant melanomas of mucous membranes. *Journal of clinical microbiology* 2008, **46**(11)**:** 3595-3598.
83. Hofbauer GF, Bavinck JNB, Euvrard S. Organ transplantation and skin cancer: basic problems and new perspectives. *Experimental dermatology* 2010, **19**(6)**:** 473-482.
84. Adami J, Gäbel H, Lindelöf B, Ekström K, Rydh B, Glimelius B*, et al.* Cancer risk following organ transplantation: a nationwide cohort study in Sweden. *British journal of cancer* 2003, **89**(7)**:** 1221-1227.
85. Kassem A, Technau K, Kurz AK, Pantulu D, Löning M, Kayser G*, et al.* Merkel cell polyomavirus sequences are frequently detected in nonmelanoma skin cancer of immunosuppressed patients. *International Journal of Cancer* 2009, **125**(2)**:** 356-361.
86. Andersson K, Michael KM, Luostarinen T, Waterboer T, Gislefoss R, Hakulinen T*, et al.* Prospective study of human papillomavirus seropositivity and risk of nonmelanoma skin cancer. *American journal of epidemiology* 2012, **175**(7)**:** 685-695.
87. Madeleine MM, Carter JJ, Johnson LG, Wipf GC, Davis C, Berg D*, et al.* Risk of squamous cell skin cancer after organ transplant associated with antibodies to cutaneous papillomaviruses, polyomaviruses, and TMC6/8 (EVER1/2) variants. *Cancer medicine* 2014.
88. Karimkhani C, Boyers LN, Dellavalle RP, Weinstock MA. It's time for" keratinocyte carcinoma" to replace the term" nonmelanoma skin cancer". *Journal of the American Academy of Dermatology* 2015, **72**(1)**:** 186.
89. Hay RJ, Johns NE, Williams HC, Bolliger IW, Dellavalle RP, Margolis DJ*, et al.* The global burden of skin disease in 2010: an analysis of the prevalence and impact of skin conditions. *Journal of Investigative Dermatology* 2014, **134**(6)**:** 1527-1534.
90. Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. *British Journal of Dermatology* 2012, **166**(5)**:** 1069-1080.
91. Karagas M, Weinstock MA, Nelson HH. Keratinocyte carcinomas (basal and squamous cell carcinomas of the skin). *Cancer epidemiology and prevention, 3rd ed. Oxford*, 2006, pp 1230-1251.
92. US Department of Health, Human Services. The surgeon general's call to action to prevent skin cancer. 2014.
93. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM*, et al. International classification of diseases for oncology*. World Health Organization, 2000.
94. Ratushny V, Gober MD, Hick R, Ridky TW, Seykora JT. From keratinocyte to cancer: the pathogenesis and modeling of cutaneous squamous cell carcinoma. *The Journal of clinical investigation* 2012, **122**(122 (2))**:** 464-472.
95. Drouin R, Therrien JP. UVB-induced cyclobutane pyrimidine dimer frequency correlates with skin cancer mutational hotspots in p53. *Photochemistry and photobiology* 1997, **66**(5)**:** 719-726.
96. Brash DE, Rudolph JA, Simon JA, Lin A, McKenna GJ, Baden HP*, et al.* A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proceedings of the National Academy of Sciences* 1991, **88**(22)**:** 10124-10128.

97. Ziegler A, Leffell DJ, Kunala S, Sharma HW, Gailani M, Simon JA*, et al.* Mutation hotspots due to sunlight in the p53 gene of nonmelanoma skin cancers. *Proceedings of the National Academy of Sciences* 1993, **90**(9)**:** 4216-4220.
98. Ayli EE, Li W, Brown TT, Witkiewicz A, Elenitsas R, Seykora JT. Activation of Src-family tyrosine kinases in hyperproliferative epidermal disorders. *Journal of cutaneous pathology* 2008, **35**(3)**:** 273-277.
99. Spencer JM, Kahn SM, Jiang W, DeLeo VA, Weinstein IB. Activated ras genes occur in human actinic keratoses, premalignant precursors to squamous cell carcinomas. *Archives of Dermatology* 1995, **131**(7)**:** 796-800.
100. Shin M, Matsuo K, Tada T, Fukushima H, Furuta H, Ozeki S*, et al.* The inhibition of RANKL/RANK signaling by osteoprotegerin suppresses bone invasion by oral squamous cell carcinoma cells. *Carcinogenesis* 2011, **32**(11)**:** 1634-1640.
101. Kwa RE, Campana K, Moy RL. Biology of cutaneous squamous cell carcinoma. *Journal of the American Academy of Dermatology* 1992, **26**(1)**:** 1-26.
102. Karia PS, Han J, Schmults CD. Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. *Journal of the American Academy of Dermatology* 2013, **68**(6)**:** 957-966.
103. Glass AG, Hoover RN. The emerging epidemic of melanoma and squamous cell skin cancer. *The Journal of the American Medical Association* 1989, **262**(15)**:** 2097-2100.
104. Karagas M, Weinstock MA, Nelson HH. Keratinocyte carcinomas (basal and squamous cell carcinomas of the skin). *Cancer epidemiology and prevention*, 3 edn. Oxford University Press, 2006, pp 1230-1250.
105. Buettner PG, Raasch BA. Incidence rates of skin cancer in Townsville, Australia. *International Journal of Cancer* 1998, **78**(5)**:** 587-593.
106. Giles GG, Marks R, Foley P. Incidence of non-melanocytic skin cancer treated in Australia. *British Medical Journal* 1988, **296**(6614)**:** 13-17.
107. Bath-Hextall F, Leonardi-Bee J, Smith C, Meal A, Hubbard R. Trends in incidence of skin basal cell carcinoma. Additional evidence from a UK primary care database study. *International Journal of Cancer* 2007, **121**(9)**:** 2105-2108.
108. Holme S, Malinovszky K, Roberts D. Changing trends in non-melanoma skin cancer in South Wales, 1988–98. *British Journal of Dermatology* 2000, **143**(6)**:** 1224-1229.
109. Carsin A, Sharp L, Comber H. Geographical, urban/rural and socioeconomic variations in nonmelanoma skin cancer incidence: a population-based study in Ireland. *British Journal of Dermatology* 2011, **164**(4)**:** 822-829.
110. Deady S, Sharp L, Comber H. Increasing skin cancer incidence in young, affluent, urban populations: a challenge for prevention. *The British journal of dermatology* 2014, **171**(2)**:** 324-331.
111. Staples MP, Elwood M, Burton RC, Williams JL, Marks R, Giles GG. Non-melanoma skin cancer in Australia: the 2002 national survey and trends since 1985. *Medical Journal of Australia* 2006, **184**(1)**:** 6-10.
112. Bielsa I, Soria X, Esteve M, Ferrandiz C. Population-based incidence of basal cell carcinoma in a Spanish Mediterranean area. *British Journal of Dermatology* 2009, **161**(6)**:** 1341-1346.
113. Boi S, Cristofolini M, Micciolo R, Polla E, Dalla Palma P. Epidemiology of skin tumors: data from the cutaneous cancer registry in Trentino, Italy. *Journal of Cutaneous Medicine and Surgery: Incorporating Medical and Surgical Dermatology* 2003, **7**(4)**:** 300-305.

114. Levi F, Franceschi S, Te V, Randimbison L, La Vecchia C. Trends of skin cancer in the Canton of Vaud, 1976-92. *British journal of cancer* 1995, **72**(4)**:** 1047.
115. Levi F, Te V, Randimbison L, Erler G, La Vecchia C. Trends in skin cancer incidence in Vaud: an update, 1976–1998. *European journal of cancer prevention* 2001, **10**(4)**:** 371-373.
116. Eisemann N, Waldmann A, Geller AC, Weinstock MA, Volkmer B, Greinert R*, et al.* Non-melanoma skin cancer incidence and impact of skin cancer screening on incidence. *Journal of Investigative Dermatology* 2014, **134**(1)**:** 43-50.
117. Dal H, Boldemann C, Lindelof B. Trends during a half century in relative squamous cell carcinoma distribution by body site in the Swedish population: support for accumulated sun exposure as the main risk factor. *The Journal of dermatology* 2008, **35**(2)**:** 55-62.
118. Lipozenčić J, Celić D, Strnad M, Tončić Jurakić R, Ledić Drvar D. Skin Cancers in Croatia. 7th World Congress on Melanoma; 5th Congress of the European Association of Dermato-Oncology (EADO); 2009; 2009.
119. Harris RB, Griffith K, Moon TE. Trends in the incidence of nonmelanoma skin cancers in southeastern Arizona, 1985-1996. *Journal of the American Academy of Dermatology* 2001, **45**(4)**:** 528-536.
120. Karagas MR, Greenberg ER, Spencer SK, Stukel TA, Mott LA. Increase in incidence rates of basal cell and squamous cell skin cancer in New Hampshire, USA. *International Journal of Cancer* 1999, **81**(4)**:** 555-559.
121. Serrano H, Scotto J, Shornick G, Fears TR, Greenberg ER. Incidence of nonmelanoma skin cancer in New Hampshire and Vermont. *Journal of the American Academy of Dermatology* 1991, **24**(4)**:** 574-579.
122. Chuang T-Y, Popescu A, Su WD, Chute CG. Basal cell carcinoma: a population-based incidence study in Rochester, Minnesota. *Journal of the American Academy of Dermatology* 1990, **22**(3)**:** 413-417.
123. Demers AA, Nugent Z, Mihalcioiu C, Wiseman MC, Kliewer EV. Trends of nonmelanoma skin cancer from 1960 through 2000 in a Canadian population. *Journal of the American Academy of Dermatology* 2005, **53**(2)**:** 320-328.
124. Jung G, Metelitsa A, Dover D, Salopek T. Trends in incidence of nonmelanoma skin cancers in Alberta, Canada, 1988–2007. *British Journal of Dermatology* 2010, **163**(1)**:** 146-154.
125. Hayes RC, Leonfellner S, Pilgrim W, Liu J, Keeling DN. Incidence of nonmelanoma skin cancer in New Brunswick, Canada, 1992 to 2001. *Journal of cutaneous medicine and surgery* 2007, **11**(2)**:** 45-52.
126. IARC. *Cancer in Africa: Epidemiology and Prevention*: Lyon, France, 2002.
127. IARC. *Cancer Incidence in Five Continents*, vol. VIII. IARC Scientific Publications: Lyon, France, 2002.
128. Koh D, Wang H, Lee J, Chia K, Lee H, Goh C. Basal cell carcinoma, squamous cell carcinoma and melanoma of the skin: analysis of the Singapore Cancer Registry data 1968–97. *British journal of Dermatology* 2003, **148**(6)**:** 1161-1166.
129. Omari AK, Khammash MR, Matalka I. Skin cancer trends in northern Jordan. *International journal of dermatology* 2006, **45**(4)**:** 384-388.
130. Sng J, Koh D, Siong WC, Choo TB. Skin cancer trends among Asians living in Singapore from 1968 to 2006. *Journal of the American Academy of Dermatology* 2009, **61**(3)**:** 426-432.

131. Nunes DH, Back L, Silva RV, Medeiros VdS. Incidence of squamous cell carcinoma of the skin in the city of Tubarao (SC)-Brazil in the years 2000, 2003, and 2006. *Anais brasileiros de dermatologia* 2009, **84**(5)**:** 482-488.
132. Scotto J, Fears TR, Fraumeni JF. Incidence of nonmelanoma skin cancer in the United States. US Department of Health and Human Services Washington; 1981.
133. Scotto J, Kopf AW, Urbach F. Non-melanoma skin cancer among caucasians in four areas of the United States. *Cancer* 1974, **34**(4)**:** 1333-1338.
134. Scotto J, Fears TR, Fraumeni JF. Incidence of nonmelanoma skin cancer in the United States. US Department of Health and Human Services Washington; 1983.
135. Qureshi AA, Laden F, Colditz GA, Hunter DJ. Geographic variation and risk of skin cancer in US women: differences between melanoma, squamous cell carcinoma, and basal cell carcinoma. *Archives of internal medicine* 2008, **168**(5)**:** 501-507.
136. Staples M, Marks R, Giles G. Trends in the incidence of non-melanocytic skin cancer (NMSC) treated in Australia 1985-1995: are primary prevention programs starting to have an effect? *International Journal of Cancer* 1998, **78**(2)**:** 144-148.
137. Karia PS, Han J, Schmults CD. Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. *Journal of the American Academy of Dermatology* 2013, **68**(6)**:** 957-966.
138. Athas WF, Hunt WC, Key CR. Changes in nonmelanoma skin cancer incidence between 1977–1978 and 1998–1999 in Northcentral New Mexico. *Cancer Epidemiology Biomarkers & Prevention* 2003, **12**(10)**:** 1105-1108.
139. Gray DT, Suman VJ, Su WD, Clay RP, Harmsen WS, Roenigk RK. Trends in the population-based incidence of squamous cell carcinoma of the skin first diagnosed between 1984 and 1992. *Archives of dermatology* 1997, **133**(6)**:** 735-740.
140. Ko C, Walton S, Keczkes K, Bury H, Nicholson C. The emerging epidemic of skin cancer. *British Journal of Dermatology* 1994, **130**(3)**:** 269-272.
141. Brewster D, Bhatti L, Inglis J, Nairn E, Doherty V. Recent trends in incidence of nonmelanoma skin cancers in the East of Scotland, 1992–2003. *British Journal of Dermatology* 2007, **156**(6)**:** 1295-1300.
142. Doherty V, Brewster D, Jensen S, Gorman D. Trends in skin cancer incidence by socioeconomic position in Scotland, 1978–2004. *British journal of cancer* 2010, **102**(11)**:** 1661-1664.
143. Robsahm TE, Helsing P, Veierød MB. Cutaneous squamous cell carcinoma in norway 1963–2011: increasing incidence and stable mortality. *Cancer medicine* 2015, **4**(3)**:** 472-480.
144. Birch-Johansen F, Jensen A, Mortensen L, Olesen AB, Kjær SK. Trends in the incidence of nonmelanoma skin cancer in Denmark 1978–2007: rapid incidence increase among young Danish women. *International journal of cancer* 2010, **127**(9)**:** 2190-2198.
145. Wassberg C, Thorn M, Johansson AM, Bergstrom R, Berne B, Ringborg U. Increasing incidence rates of squamous cell carcinoma of the skin in Sweden. *Acta Dermato-Venereologica* 2001, **81**(4)**:** 268-272.
146. Hannuksela-Svahn A, Pukkala E, Karvonen J. Basal cell skin carcinoma and other nonmelanoma skin cancers in Finland from 1956 through 1995. *Archives of dermatology* 1999, **135**(7)**:** 781-786.

147. Holterhues C, de Vries E, Louwman MW, Koljenović S, Nijsten T. Incidence and trends of cutaneous malignancies in the Netherlands, 1989–2005. *Journal of Investigative Dermatology* 2010, **130**(7)**:** 1807-1812.
148. Flohil SC, Seubring I, van Rossum MM, Coebergh J-WW, de Vries E, Nijsten T. Trends in Basal Cell Carcinoma Incidence Rates: A 37-Year Dutch Observational Study. *Journal of Investigative Dermatology* 2013, **133**(4)**:** 913-918.
149. Hollestein L, de Vries E, Nijsten T. Trends of cutaneous squamous cell carcinoma in the Netherlands: increased incidence rates, but stable relative survival and mortality 1989–2008. *European Journal of Cancer* 2012, **48**(13)**:** 2046-2053.
150. Hollestein LM, de Vries E, Aarts MJ, Schroten C, Nijsten TE. Burden of disease caused by keratinocyte cancer has increased in The Netherlands since 1989. *Journal of the American Academy of Dermatology* 2014, **71**(5)**:** 896-903.
151. Flohil SC, De Vries E, Neumann M, COEBERgH J-W, Nijsten T. Incidence, prevalence and future trends of primary basal cell carcinoma in the Netherlands. *Acta dermato-venereologica* 2011, **91**(1)**:** 24-30.
152. Perera E, Gnaneswaran N, Staines C, Win AK, Sinclair R. Incidence and prevalence of non-melanoma skin cancer in Australia: A systematic review. *The Australasian journal of dermatology* 2015.
153. Videnović G, Miljuš D, Ilić D, Krasić D, Živković S. Nonmelanoma skin cancer in the population of the city of Belgrade in the period 1999-2011. *Srpski arhiv za celokupno lekarstvo* 2015, **143**(5-6)**:** 290-295.
154. Razi S, Rafiemanesh H, Ghoncheh M, Khani Y, Salehiniya H. Changing Trends of Types of Skin Cancer in Iran. *Asian Pacific journal of cancer prevention: APJCP* 2014, **16**(12)**:** 4955-4958.
155. Slaper H, Velders GJ, Daniel JS, de Gruijl FR, van der Leun JC. Estimates of ozone depletion and skin cancer incidence to examine the Vienna Convention achievements. *Nature* 1996, **384**(6606)**:** 256-258.
156. Lucas RM, Norval M, Neale RE, Young AR, de Gruijl FR, Takizawa Y*, et al.* The consequences for human health of stratospheric ozone depletion in association with other environmental factors. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* 2015, **14**(1)**:** 53-87.
157. Christenson LJ, Borrowman TA, Vachon CM, Tollefson MM, Otley CC, Weaver AL*, et al.* Incidence of basal cell and squamous cell carcinomas in a population younger than 40 years. *The Journal of the American Medical Association* 2005, **294**(6)**:** 681-690.
158. Valery PC, Neale R, Williams G, Pandeya N, Siller G, Green A. The effect of skin examination surveys on the incidence of basal cell carcinoma in a Queensland community sample: a 10-year longitudinal study. Journal of Investigative Dermatology Symposium Proceedings; 2004: Nature Publishing Group; 2004. p. 148-151.
159. Greenberg ER, Baron JA, Stukel TA, Stevens MM, Mandel JS, Spencer SK*, et al.* A clinical trial of beta carotene to prevent basal-cell and squamous-cell cancers of the skin. *New England Journal of Medicine* 1990, **323**(12)**:** 789-795.
160. Wechsler J, Zanetti R, Schrameck C, Rosso S, Pippione M, Linares J*, et al.* Reproducibility of histopathologic diagnosis and classification of non-melanocytic skin cancer: a panel exercise in the framework of the multicenter southern European study HELIOS. *Tumori* 2000, **87**(2)**:** 95-100.

161. Davis DA, Donahue JP, Bost JE, Horn TD. The diagnostic concordance of actinic keratosis and squamous cell carcinoma. *Journal of cutaneous pathology* 2005, **32**(1): 84-84.
162. Karjalainen S, Salo H, Teppo L. Basal cell and squamous cell carcinoma of the skin in Finland. Site distribution and patient survival. *International journal of dermatology* 1989, **28**(7): 445-450.
163. Schmults CD, Karia PS, Carter JB, Han J, Qureshi AA. Factors predictive of recurrence and death from cutaneous squamous cell carcinoma: a 10-year, single-institution cohort study. *JAMA dermatology* 2013, **149**(5): 541-547.
164. Alam M, Ratner D. Cutaneous squamous-cell carcinoma. *New England Journal of Medicine* 2001, **344**(13): 975-983.
165. Jensen AO, Lamberg AL, Jacobsen JB, Braae Olesen A, Sorensen HT. Non-melanoma skin cancer and ten-year all-cause mortality: a population-based cohort study. *Acta Dermato Venereologica* 2010, **90**(4): 362-367.
166. Jambusaria-Pahlajani A, Kanetsky PA, Karia PS, Hwang W-T, Gelfand JM, Whalen FM, *et al.* Evaluation of AJCC tumor staging for cutaneous squamous cell carcinoma and a proposed alternative tumor staging system. *JAMA dermatology* 2013, **149**(4): 402-410.
167. Rees JR, Zens MS, Celaya MO, Riddle BL, Karagas MR, Peacock JL. Survival after squamous cell and basal cell carcinoma of the skin: A retrospective cohort analysis. *International Journal of Cancer* 2015, **137**(4): 878-884.
168. Jensen AØ, Bautz A, Olesen AB, Karagas M, Sørensen HT, Friis S. Mortality in Danish patients with nonmelanoma skin cancer, 1978–2001. *British Journal of Dermatology* 2008, **159**(2): 419-425.
169. Jensen AO, Olesen AB, Dethlefsen C, Sorensen HT. Ten year mortality in a cohort of nonmelanoma skin cancer patients in Denmark. *Journal of Investigative Dermatology* 2006, **126**(11): 2539-2541.
170. English DR, Armstrong BK, Kricker A, Winter MG, Heenan PJ, Randell PL. Case-control study of sun exposure and squamous cell carcinoma of the skin. *International Journal of Cancer* 1998, **77**(3): 347-353.
171. Gallagher RP, Hill GB, Bajdik CD, Coldman AJ, Fincham S, McLean DI, *et al.* Sunlight exposure, pigmentation factors, and risk of nonmelanocytic skin cancer: II. Squamous cell carcinoma. *Archives of Dermatology* 1995, **131**(2): 164-169.
172. Elwood J, Gallagher R, Hill G, Spinelli J, Pearson J, Threlfall W. Pigmentation and skin reaction to sun as risk factors for cutaneous melanoma: Western Canada Melanoma Study. *British medical journal (Clinical research ed)* 1984, **288**(6411): 99.
173. Lichter MD, Karagas MR, Mott LA, Spencer SK, Stukel TA, Greenberg ER. Therapeutic ionizing radiation and the incidence of basal cell carcinoma and squamous cell carcinoma. *Archives of Dermatology* 2000, **136**(8): 1007-1011.
174. Karagas MR, McDonald JA, Greendberg ER, Stukel TA, Weiss JE, Baron JA, *et al.* Risk of basal cell and squamous cell skin cancers after ionizing radiation therapy. *Journal of the National Cancer Institute* 1996, **88**(24): 1848-1853.
175. Gilbert-Diamond D, Li Z, Perry AE, Spencer SK, Gandolfi AJ, Karagas MR. A population-based case–control study of urinary arsenic species and squamous cell carcinoma in New Hampshire, USA. *Environmental health perspectives* 2013, **121**(10): 1154.

176. Chen Y-C, Guo Y-LL, Su H-JJ, Hsueh Y-M, Smith TJ, Ryan LM*, et al.* Arsenic methylation and skin cancer risk in southwestern Taiwan. *Journal of occupational and environmental medicine* 2003, **45**(3)**:** 241-248.
177. Karagas MR, Stukel TA, Tosteson TD. Assessment of cancer risk and environmental levels of arsenic in New Hampshire. *International journal of hygiene and environmental health* 2002, **205**(1)**:** 85-94.
178. Karagas MR, Stukel TA, Morris JS, Tosteson TD, Weiss JE, Spencer SK*, et al.* Skin cancer risk in relation to toenail arsenic concentrations in a US population-based case-control study. *American Journal of Epidemiology* 2001, **153**(6)**:** 559-565.
179. Everall J, Dowd P. Influence of environmental factors excluding ultra violet radiation on the incidence of skin cancer. *Bulletin du cancer* 1977, **65**(3)**:** 241-247.
180. Siddens LK, Larkin A, Krueger SK, Bradfield CA, Waters KM, Tilton SC*, et al.* Polycyclic aromatic hydrocarbons as skin carcinogens: Comparison of benzo[a]pyrene, dibenzo[def,p]chrysene and three environmental mixtures in the FVB/N mouse. *Toxicology and applied pharmacology* 2012, **264**(3)**:** 377-386.
181. IARC. *Exposure to artificial UV radiation and skin cancer*, vol. 1. World Health Organization, 2006.
182. Lazovich D, Vogel RI, Berwick M, Weinstock MA, Anderson KE, Warshaw EM. Indoor tanning and risk of melanoma: a case-control study in a highly exposed population. *Cancer Epidemiology Biomarkers & Prevention* 2010, **19**(6)**:** 1557-1568.
183. Zhang M, Qureshi AA, Geller AC, Frazier L, Hunter DJ, Han J. Use of tanning beds and incidence of skin cancer. *Journal of Clinical Oncology* 2012, **30**(14)**:** 1588-1593.
184. Wehner MR, Chren M-M, Nameth D, Choudhry A, Gaskins M, Nead KT*, et al.* International prevalence of indoor tanning: a systematic review and meta-analysis. *JAMA dermatology* 2014, **150**(4)**:** 390-400.
185. Lautenschlager S, Wulf HC, Pittelkow MR. Photoprotection. *The Lancet* 2007, **370**(9586)**:** 528-537.
186. Pouplard C, Brenaut E, Horreau C, Barnetche T, Misery L, Richard MA*, et al.* Risk of cancer in psoriasis: a systematic review and meta-analysis of epidemiological studies. *Journal of the European Academy of Dermatology and Venereology* 2013, **27 Suppl 3:** 36-46.
187. Ji J, Shu X, Sundquist K, Sundquist J, Hemminki K. Cancer risk in hospitalised psoriasis patients: a follow-up study in Sweden. *British journal of cancer* 2009, **100**(9)**:** 1499-1502.
188. Stern RS. The risk of squamous cell and basal cell cancer associated with psoralen and ultraviolet A therapy: a 30-year prospective study. *Journal of the American Academy of Dermatology* 2012, **66**(4)**:** 553-562.
189. Armstrong BK, Kricker A, English DR. Sun exposure and skin cancer. *Australasian Journal of Dermatology* 1997, **38**(S1)**:** S1-S6.
190. Brash DE. Sunlight and the onset of skin cancer. *Trends in genetics* 1997, **13**(10)**:** 410-414.
191. English DR, Armstrong BK, Kricker A, Fleming C. Sunlight and cancer. *Cancer causes & control* 1997, **8**(3)**:** 271-283.
192. Welsh MM, Karagas MR, Applebaum KM, Spencer SK, Perry AE, Nelson HH. A role for ultraviolet radiation immunosuppression in non-melanoma skin cancer as evidenced by gene-environment interactions. *Carcinogenesis* 2008, **29**(10)**:** 1950-1954.

193. Warren R, Gartstein V, Kligman AM, Montagna W, Allendorf RA, Ridder GM. Age, sunlight, and facial skin: a histologic and quantitative study. *Journal of the American Academy of Dermatology* 1991, **25**(5)**:** 751-760.
194. Battistutta D, Pandeya N, Strutton GM, Fourtanier A, Tison S, Green AC. Skin surface topography grading is a valid measure of skin photoaging. *Photodermatology, photoimmunology & photomedicine* 2006, **22**(1)**:** 39-45.
195. Lucas RM, Ponsonby A-L, Dear K, Taylor BV, Dwyer T, McMichael AJ*, et al.* Associations between silicone skin cast score, cumulative sun exposure, and other factors in the ausimmune study: a multicenter Australian study. *Cancer Epidemiology Biomarkers & Prevention* 2009, **18**(11)**:** 2887-2894.
196. Green A, Hughes M, McBride P, Fourtanier A. Factors associated with premature skin aging (photoaging) before the age of 55: a population-based study. *Dermatology* 2011, **222**(1)**:** 74-80.
197. Hughes MC, Strutton GM, Fourtanier A, Green AC. Validation of skin surface microtopography as a measure of skin photoaging in a subtropical population aged 40 and over. *Photodermatology, photoimmunology & photomedicine* 2012, **28**(3)**:** 153-158.
198. De Vries E, Trakatelli M, Kalabalikis D, Ferrandiz L, Ruiz-de-Casas A, Moreno-Ramirez D*, et al.* Known and potential new risk factors for skin cancer in European populations: a multicentre case-control study. *The British journal of dermatology* 2012, **167 Suppl 2:** 1-13.
199. Rizzato C, Scherer D, Rudnai P, Gurzau E, Koppova K, Hemminki K*, et al.* POMC and TP53 genetic variability and risk of basal cell carcinoma of skin: Interaction between host and genetic factors. *Journal of dermatological science* 2011, **63**(1)**:** 47-54.
200. Ferrucci LM, Cartmel B, Molinaro AM, Gordon PB, Leffell DJ, Bale AE*, et al.* Host phenotype characteristics and MC1R in relation to early-onset basal cell carcinoma. *Journal of Investigative Dermatology* 2012, **132**(4)**:** 1272-1279.
201. Kricker A, Armstrong BK, English DR, Heenan PJ. Pigmentary and cutaneous risk factors for non-melanocytic skin cancer—A case-control study. *International journal of cancer* 1991, **48**(5)**:** 650-662.
202. English DR, Armstrong BK, Kricker A, Winter MG, Heenan PJ, Randell PL. Demographic characteristics, pigmentary and cutaneous risk factors for squamous cell carcinoma of the skin: A case-control study. *International Journal of Cancer* 1998, **76**(5)**:** 628-634.
203. Veierod MB, Couto E, Lund E, Adami HO, Weiderpass E. Host characteristics, sun exposure, indoor tanning and risk of squamous cell carcinoma of the skin. *International journal of cancer* 2014, **135**(2)**:** 413-422.
204. Suarez B, Lopez-Abente G, Martinez C, Navarro C, Tormo MJ, Rosso S*, et al.* Occupation and skin cancer: the results of the HELIOS-I multicenter case-control study. *BMC public health* 2007, **7:** 180.
205. Yoshinaga S, Hauptmann M, Sigurdson AJ, Doody MM, Freedman DM, Alexander BH*, et al.* Nonmelanoma skin cancer in relation to ionizing radiation exposure among US radiologic technologists. *International journal of cancer* 2005, **115**(5)**:** 828-834.
206. Kitahara CM, Linet MS, Drozdovitch V, Alexander BH, Preston DL, Simon SL*, et al.* Cancer and circulatory disease risks in US radiologic technologists associated with performing procedures involving radionuclides. *Occupational and environmental medicine* 2015.

207. Cherrie JW, Semple S. Dermal exposure to metalworking fluids and medium-chain chlorinated paraffin (MCCP). *Annals of occupational hygiene* 2010, **54**(2)**:** 228-235.
208. Karagas MR. Occurrence of cutaneous basal cell and squamous cell malignancies among those with a prior history of skin cancer. The Skin Cancer Prevention Study Group. *The Journal of investigative dermatology* 1994, **102**(6)**:** 10S-13S.
209. Karagas MR, Stukel TA, Greenberg E, Baron JA, Mott LA, Stern RS. RIsk of subsequent basal cell carcinoma and squamous cell carcinoma of the skin among patients with prior skin cancer. *JAMA* 1992, **267**(24)**:** 3305-3310.
210. Flohil SC, van der Leest RJ, Arends LR, de Vries E, Nijsten T. Risk of subsequent cutaneous malignancy in patients with prior keratinocyte carcinoma: a systematic review and meta-analysis. *European Journal of Cancer* 2013, **49**(10)**:** 2365-2375.
211. Dumaz N, Drougard C, Sarasin A, Daya-Grosjean L. Specific UV-induced mutation spectrum in the p53 gene of skin tumors from DNA-repair-deficient xeroderma pigmentosum patients. *Proceedings of the National Academy of Sciences* 1993, **90**(22)**:** 10529-10533.
212. Van Steeg H, Kraemer KH. Xeroderma pigmentosum and the role of UV-induced DNA damage in skin cancer. *Molecular medicine today* 1999, **5**(2)**:** 86-94.
213. Birkeland SA, Storm HH, Lamm LU, Barlow L, Blohmé I, Forsberg B*, et al.* Cancer risk after renal transplantation in the Nordic countries, 1964–1986. *International journal of cancer* 1995, **60**(2)**:** 183-189.
214. Kinlen L, Sheil A, Peto J, Doll R. Collaborative United Kingdom-Australasian study of cancer in patients treated with immunosuppressive drugs. *British medical journal* 1979, **2**(6203)**:** 1461.
215. Anforth R, Menzies A, Byth K, Carlos G, Chou S, Sharma R*, et al.* Factors influencing the development of cutaneous squamous cell carcinoma in patients on BRAF inhibitor therapy. *Journal of the American Academy of Dermatology* 2015, **72**(5)**:** 809-815.e801.
216. Robinson SN, Zens MS, Perry AE, Spencer SK, Duell EJ, Karagas MR. Photosensitizing agents and the risk of non-melanoma skin cancer: a population-based case-control study. *Journal of Investigative Dermatology* 2013, **133**(8)**:** 1950-1955.
217. Karagas MR, Stukel TA, Umland V, Tsoukas MM, Mott LA, Sorensen HT*, et al.* Reported use of photosensitizing medications and basal cell and squamous cell carcinoma of the skin: results of a population-based case-control study. *Journal of Investigative Dermatology* 2007, **127**(12)**:** 2901-2903.
218. Kaae J, Boyd HA, Hansen AV, Wulf HC, Wohlfahrt J, Melbye M. Photosensitizing medication use and risk of skin cancer. *Cancer Epidemiology, Biomarkers & Prevention* 2010, **19**(11)**:** 2942-2949.
219. Jensen AØ, Thomsen H, Engebjerg MC, Olesen AB, Sørensen HT, Karagas M. Use of photosensitising diuretics and risk of skin cancer: a population-based case–control study. *British journal of cancer* 2008, **99**(9)**:** 1522-1528.
220. Schmidt SA, Schmidt M, Mehnert F, Lemeshow S, Sorensen HT. Use of antihypertensive drugs and risk of skin cancer. *Journal of the European Academy of Dermatology and Venereology: JEADV* 2015, **29**(8)**:** 1545-1554.
221. Vessey M, Painter R, Powell J. Skin disorders in relation to oral contraception and other factors, including age, social class, smoking and body mass index. Findings in a large cohort study. *British Journal of Dermatology* 2000, **143**(4)**:** 815-820.

222. Applebaum KM, Nelson HH, Zens MS, Stukel TA, Spencer SK, Karagas MR. Oral Contraceptives: A Risk Factor for Squamous Cell Carcinoma&quest. *Journal of Investigative Dermatology* 2009, **129**(12)**:** 2760-2765.
223. Asgari MM, Efird JT, Warton EM, Friedman GD. Potential risk factors for cutaneous squamous cell carcinoma include oral contraceptives: results of a nested case-control study. *International journal of environmental research and public health* 2010, **7**(2)**:** 427-442.
224. Birch-Johansen F, Jensen A, Olesen AB, Christensen J, Tjønneland A, Kjær SK. Does hormone replacement therapy and use of oral contraceptives increase the risk of non-melanoma skin cancer? *Cancer Causes & Control* 2012, **23**(2)**:** 379-388.
225. Cahoon EK, Kitahara CM, Ntowe E, Bowen EM, Doody MM, Alexander BH*, et al.* Female Estrogen-Related Factors and Incidence of Basal Cell Carcinoma in a Nationwide US Cohort. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 2015, **33**(34)**:** 4058-4065.
226. Efird JT, Toland AE, Lea CS, Phillips CJ. The combined influence of oral contraceptives and human papillomavirus virus on cutaneous squamous cell carcinoma. *Clinical Medicine Insights Oncology* 2011, **5:** 55.
227. Tang JY, Spaunhurst KM, Chlebowski RT, Wactawski-Wende J, Keiser E, Thomas F*, et al.* Menopausal hormone therapy and risks of melanoma and nonmelanoma skin cancers: women's health initiative randomized trials. *Journal of the National Cancer Institute* 2011, **103**(19)**:** 1469-1475.
228. Johannesdottir SA, Chang ET, Mehnert F, Schmidt M, Olesen AB, Sorensen HT. Nonsteroidal anti-inflammatory drugs and the risk of skin cancer: a population-based case-control study. *Cancer* 2012, **118**(19)**:** 4768-4776.
229. Clouser MC, Roe DJ, Foote JA, Harris RB. Effect of non-steroidal anti-inflammatory drugs on non-melanoma skin cancer incidence in the SKICAP-AK trial. *Pharmacoepidemiology and drug safety* 2009, **18**(4)**:** 276-283.
230. Butler GJ, Neale R, Green AC, Pandeya N, Whiteman DC. Nonsteroidal anti-inflammatory drugs and the risk of actinic keratoses and squamous cell cancers of the skin. *Journal of the American Academy of Dermatology* 2005, **53**(6)**:** 966-972.
231. Zhang B, Liang X, Ye L, Wang Y. No chemopreventive effect of nonsteroidal anti-inflammatory drugs on nonmelanoma skin cancer: Evidence from meta-analysis. 2014.
232. Muranushi C, Olsen CM, Pandeya N, Green AC. Aspirin and nonsteroidal anti-inflammatory drugs can prevent cutaneous squamous cell carcinoma: a systematic review and meta-analysis. *Journal of Investigative Dermatology* 2015, **135**(4)**:** 975-983.
233. Cheng J, Zens MS, Duell E, Perry AE, Chapman MS, Karagas MR. History of Allergy and Atopic Dermatitis in Relation to Squamous Cell and Basal Cell Carcinoma of the Skin. *Cancer Epidemiology Biomarkers & Prevention* 2015, **24**(4)**:** 749-754.
234. Eriksson N, Mikoczy Z, Hagmar L. Cancer incidence in 13,811 patients skin tested for allergy. *Journal of Investigational Allergology and Clinical Immunology* 2005, **15**(3)**:** 161.
235. Ji J, Shu X, Li X, Sundquist K, Sundquist J, Hemminki K. Cancer risk in hospitalised asthma patients. *British journal of cancer* 2009, **100**(5)**:** 829-833.
236. Jensen AO, Svaerke C, Farkas DK, Olesen AB, Kragballe K, Sørensen HT. Atopic Dermatitis and Risk of Skin Cancer. *American journal of clinical dermatology* 2012, **13**(1)**:** 29-36.

237. Milan T, Verkasalo P, Kaprio J, Koskenvuo M. Lifestyle differences in twin pairs discordant for basal cell carcinoma of the skin. *British Journal of Dermatology* 2003, **149**(1)**:** 115-123.
238. Ibiebele TI, van der Pols JC, Hughes MC, Marks GC, Green AC. Dietary fat intake and risk of skin cancer: a prospective study in Australian adults. *International journal of cancer* 2009, **125**(7)**:** 1678-1684.
239. Hughes MC, van der Pols JC, Marks GC, Green AC. Food intake and risk of squamous cell carcinoma of the skin in a community: the Nambour skin cancer cohort study. *International journal of cancer* 2006, **119**(8)**:** 1953-1960.
240. Ibiebele TI, van der Pols JC, Hughes MC, Marks GC, Williams GM, Green AC. Dietary pattern in association with squamous cell carcinoma of the skin: a prospective study. *American Journal of Clinical Nutrition* 2007, **85**(5)**:** 1401-1408.
241. Noel SE, Stoneham AC, Olsen CM, Rhodes LE, Green AC. Consumption of omega-3 fatty acids and the risk of skin cancers: a systematic review and meta-analysis. *International journal of cancer* 2014, **135**(1)**:** 149-156.
242. Wallingford SC, van As JA, Hughes MC, Ibiebele TI, Green AC, van der Pols JC. Intake of omega-3 and omega-6 fatty acids and risk of basal and squamous cell carcinomas of the skin: a longitudinal community-based study in Australian adults. *Nutrition and cancer* 2012, **64**(7)**:** 982-990.
243. Harris RB, Foote JA, Hakim IA, Bronson DL, Alberts DS. Fatty acid composition of red blood cell membranes and risk of squamous cell carcinoma of the skin. *Cancer Epidemiology Biomarkers & Prevention* 2005, **14**(4)**:** 906-912.
244. Hakim IA, Harris RB, Weisgerber UM. Tea intake and squamous cell carcinoma of the skin: influence of type of tea beverages. *Cancer Epidemiology Biomarkers & Prevention* 2000, **9**(7)**:** 727-731.
245. Wallingford SC, Hughes MC, Green AC, van der Pols JC. Plasma omega-3 and omega-6 concentrations and risk of cutaneous basal and squamous cell carcinomas in Australian adults. *Cancer Epidemiology Biomarkers & Prevention* 2013.
246. Black HS, Thornby JI, Wolf JE, Jr., Goldberg LH, Herd JA, Rosen T*, et al.* Evidence that a low-fat diet reduces the occurrence of non-melanoma skin cancer. *International journal of cancer* 1995, **62**(2)**:** 165-169.
247. Kune GA, Bannerman S, Field B, Watson LF, Cleland H, Merenstein D*, et al.* Diet, alcohol, smoking, serum beta-carotene, and vitamin A in male nonmelanocytic skin cancer patients and controls. *Nutrition and cancer* 1992, **18**(3)**:** 237-244.
248. Jensen A, Birch-Johansen F, Olesen AB, Christensen J, Tjonneland A, Kjaer SK. Intake of alcohol may modify the risk for non-melanoma skin cancer: results of a large Danish prospective cohort study. *Journal of Investigative Dermatology* 2012, **132**(12)**:** 2718-2726.
249. Ansems TM, van der Pols JC, Hughes MC, Ibiebele T, Marks GC, Green AC. Alcohol intake and risk of skin cancer: a prospective study. *European journal of clinical nutrition* 2008, **62**(2)**:** 162-170.
250. Asgari MM, White E, Warton EM, Hararah MK, Friedman GD, Chren MM. Association of tea consumption and cutaneous squamous cell carcinoma. *Nutrition and cancer* 2011, **63**(2)**:** 314-318.

251. Rees JR, Stukel TA, Perry AE, Zens MS, Spencer SK, Karagas MR. Tea consumption and basal cell and squamous cell skin cancer: results of a case-control study. *Journal of the American Academy of Dermatology* 2007, **56**(5)**:** 781-785.
252. Ferrucci LM, Cartmel B, Molinaro AM, Leffell DJ, Bale AE, Mayne ST. Tea, coffee, and caffeine and early-onset basal cell carcinoma in a case-control study. *European Journal of Cancer Prevention* 2014, **23**(4)**:** 296-302.
253. Hakim IA, Harris RB. Joint effects of citrus peel use and black tea intake on the risk of squamous cell carcinoma of the skin. *BMC Dermatology* 2001, **1**(1)**:** 3.
254. Abel EL, Hendrix SO, McNeeley SG, Johnson KC, Rosenberg CA, Mossavar-Rahmani Y*, et al.* Daily coffee consumption and prevalence of nonmelanoma skin cancer in Caucasian women. *European Journal of Cancer Prevention* 2007, **16**(5)**:** 446-452.
255. Miura K, Hughes MC, Green AC, van der Pols JC. Caffeine intake and risk of basal cell and squamous cell carcinomas of the skin in an 11-year prospective study. *European journal of nutrition* 2014, **53**(2)**:** 511-520.
256. Song F, Qureshi AA, Han J. Increased caffeine intake is associated with reduced risk of basal cell carcinoma of the skin. *Cancer research* 2012, **72**(13)**:** 3282-3289.
257. Carless MA, Kraska T, Lintell N, Neale RE, Green AC, Griffiths LR. Polymorphisms of the VDR gene are associated with presence of solar keratoses on the skin. *The British journal of dermatology* 2008, **159**(4)**:** 804-810.
258. van der Pols JC, Russell A, Bauer U, Neale RE, Kimlin MG, Green AC. Vitamin D status and skin cancer risk independent of time outdoors: 11-year prospective study in an Australian community. *Journal of Investigative Dermatology* 2013, **133**(3)**:** 637-641.
259. Moloney F, Vestergaard M, Radojkovic B, Damian D. Randomized, double-blinded, placebo controlled study to assess the effect of topical 1% nicotinamide on actinic keratoses. *British Journal of Dermatology* 2010, **162**(5)**:** 1138-1139.
260. Chen AC, Martin AJ, Choy B, Fernández-Peñas P, Dalziell RA, McKenzie CA*, et al.* A Phase 3 Randomized Trial of Nicotinamide for Skin-Cancer Chemoprevention. *New England Journal of Medicine* 2015, **373**(17)**:** 1618-1626.
261. Bavinck J, Tieben LM, Van Der Woude FJ, Tegzess AM, Hermans J, Ter Schegget J*, et al.* Prevention of skin cancer and reduction of keratotic skin lesions during acitretin therapy in renal transplant recipients: a double-blind, placebo-controlled study. *Journal of Clinical Oncology* 1995, **13**(8)**:** 1933-1938.
262. George R, Weightman W, Russ GR, Bannister KM, Mathew TH. Acitretin for chemoprevention of non-melanoma skin cancers in renal transplant recipients. *Australasian journal of dermatology* 2002, **43**(4)**:** 269-273.
263. Moon TE, Levine N, Cartmel B, Bangert JL. Retinoids in prevention of skin cancer. *Cancer letters* 1997, **114**(1)**:** 203-205.
264. Nijsten TE, Stern RS. Oral retinoid use reduces cutaneous squamous cell carcinoma risk in patients with psoriasis treated with psoralen-UVA: a nested cohort study. *Journal of the American Academy of Dermatology* 2003, **49**(4)**:** 644-650.
265. McKenna D, Murphy G. Skin cancer chemoprophylaxis in renal transplant recipients: 5 years of experience using low-dose acitretin. *The British journal of dermatology* 1999, **140**(4)**:** 656-660.
266. Johnson JD, Houchens DP, Kluwe WM, Craig DK, Fisher GL. Effects of mainstream and environmental tobacco smoke on the immune system in animals and humans: a review. *CRC Critical Reviews in Toxicology* 1990, **20**(5)**:** 369-395.

267. Scharffetter-Kochanek K, Wlaschek M, Brenneisen P, Schauen M, Blaudschun R, Wenk J. UV-induced reactive oxygen species in photocarcinogenesis and photoaging. *Biological chemistry* 1997, **378**(11)**:** 1247-1258.
268. Kennedy C, Bajdik CD, Willemze R, de Gruijl FR, Bavinck JNB. The influence of painful sunburns and lifetime sun exposure on the risk of actinic keratoses, seborrheic warts, melanocytic nevi, atypical nevi, and skin cancer. *Journal of Investigative Dermatology* 2003, **120**(6)**:** 1087-1093.
269. Morita A. Tobacco smoke causes premature skin aging. *Journal of dermatological science* 2007, **48**(3)**:** 169-175.
270. Goldenberg A, Ortiz A, Kim SS, Jiang SB. Squamous cell carcinoma with aggressive subclinical extension: 5-year retrospective review of diagnostic predictors. *Journal of the American Academy of Dermatology* 2015, **73**(1)**:** 120-126.
271. Odenbro A, Bellocco R, Boffetta P, Lindelof B, Adami J. Tobacco smoking, snuff dipping and the risk of cutaneous squamous cell carcinoma: a nationwide cohort study in Sweden. *British journal of cancer* 2005, **92**(7)**:** 1326-1328.
272. Sanchez G, Nova J. Risk factors for squamous cell carcinoma, a study by the National Dermatology Centre of Colombia. *Actas dermo-sifiliograficas* 2013, **104**(8)**:** 672-678.
273. Rollison DE, Iannacone MR, Messina JL, Glass LF, Giuliano AR, Roetzheim RG*, et al.* Case-control study of smoking and non-melanoma skin cancer. *Cancer causes & control : CCC* 2012, **23**(2)**:** 245-254.
274. McBride P, Olsen CM, Green AC. Tobacco smoking and cutaneous squamous cell carcinoma: a 16-year longitudinal population-based study. *Cancer Epidemiology, Biomarkers & Prevention* 2011, **20**(8)**:** 1778-1783.
275. Euvrard S, Kanitakis J, Claudy A. Skin cancers after organ transplantation. *New England Journal of Medicine* 2003, **348**(17)**:** 1681-1691.
276. Pass F, Reissig M, Shah KV, Eisinger M, Orth G. Identification of an immunologically distinct papillomavirus from lesions of epidermodysplasia verruciformis. *Journal of the National Cancer Institute* 1977, **59**(4)**:** 1107-1112.
277. Schottenfeld D, Fraumeni J. Cancer epidemiology and prevention. 3rd edn Oxford University Press. Oxford; 2006.
278. Karagas MR, Waterboer T, Li Z, Nelson HH, Michael KM, Bavinck JNB*, et al.* Genus β human papillomaviruses and incidence of basal cell and squamous cell carcinomas of skin: population based case-control study. *British Medical Journal* 2010, **341**.
279. Madeleine MM, Carter JJ, Johnson LG, Wipf GC, Davis C, Berg D*, et al.* Risk of squamous cell skin cancer after organ transplant associated with antibodies to cutaneous papillomaviruses, polyomaviruses, and TMC6/8 (EVER1/2) variants. *Cancer medicine* 2014, **3**(5)**:** 1440-1447.
280. Kitamura T, Kunitake T, Guo J, Tominaga T, Kawabe K, Yogo Y. Transmission of the human polyomavirus JC virus occurs both within the family and outside the family. *Journal of clinical microbiology* 1994, **32**(10)**:** 2359-2363.
281. Teras LR, Rollison DE, Pawlita M, Michel A, Blase JL, Willhauck-Fleckenstein M*, et al.* Prediagnostic Circulating Polyomavirus Antibody Levels and Risk of non-Hodgkin Lymphoma. *Cancer Epidemiology Biomarkers & Prevention* 2014**:** cebp. 1125.2014.
282. van der Meijden E, Bialasiewicz S, Rockett RJ, Tozer SJ, Sloots TP, Feltkamp MC. Different Serologic Behavior of MCPyV, TSPyV, HPyV6, HPyV7 and HPyV9 Polyomaviruses Found on the Skin. *PloS One* 2013, **8**(11)**:** e81078.

283. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004, **5**(1): 113.
284. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. *American journal of epidemiology* 1993, **138**(11): 923-936.
285. Marsit CJ, Koestler DC, Christensen BC, Karagas MR, Houseman EA, Kelsey KT. DNA methylation array analysis identifies profiles of blood-derived DNA methylation associated with bladder cancer. *Journal of Clinical Oncology* 2011, **29**(9): 1133-1139.
286. Koestler DC, Marsit CJ, Christensen BC, Accomando W, Langevin SM, Houseman EA, *et al.* Peripheral blood immune cell methylation profiles are associated with nonhematopoietic cancers. *Cancer Epidemiology Biomarkers & Prevention* 2012, **21**(8): 1293-1302.
287. Koestler DC, Avissar-Whiting M, Houseman EA, Karagas MR, Marsit CJ. Differential DNA Methylation in Umbilical Cord Blood of Infants Exposed to Low Levels of Arsenic in Utero. *Environmental health perspectives* 2013.
288. Bock C. Analysing and interpreting DNA methylation data. *Nature Reviews Genetics* 2012, **13**(10): 705-719.
289. Du P, Zhang X, Huang C-C, Jafari N, Kibbe W, Hou L, *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010, **11**(1): 587.
290. Banister CE, Koestler DC, Maccani MA, Padbury JF, Houseman EA, Marsit CJ. Infant growth restriction is associated with distinct patterns of DNA methylation in human placentas. *Epigenetics* 2011, **6**(7): 920-927.
291. Chen Y-a, Choufani S, Ferreira JC, Grafodatskaya D, Butcher DT, Weksberg R. Sequence overlap between autosomal and sex-linked probes on the Illumina HumanMethylation27 microarray. *Genomics* 2011, **97**(4): 214-222.
292. Chen Y-a, Choufani S, Grafodatskaya D, Butcher DT, Ferreira JC, Weksberg R. Cross-reactive DNA microarray probes lead to false discovery of autosomal sex-associated DNA methylation. *American journal of human genetics* 2012, **91**(4): 762.
293. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012, **13**(1): 86.
294. Storey JD. A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2002, **64**(3): 479-498.
295. Storey JD, Taylor JE, Siegmund D. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2004, **66**(1): 187-205.
296. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences* 2003, **100**(16): 9440-9445.
297. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, *et al.* Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genetics* 2009, **5**(8): e1000602.
298. Liu J, Morgan M, Hutchison K, Calhoun VD. A study of the influence of sex on genome wide methylation. *PloS One* 2010, **5**(4): e10028.
299. Lowe R, Rakyan VK. Correcting for cell-type composition bias in epigenome-wide association studies. *Genome Medicine* 2014, **6**(3): 23.

300. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biology* 2014, **15**(2): R31.
301. Da Wei Huang BTS, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 2008, **4**(1): 44-57.
302. Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, *et al.* DAVID: database for annotation, visualization, and integrated discovery. *Genome Biology* 2003, **4**(5): P3.
303. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 2009, **4**(1): 44-57.
304. Geeleher P, Hartnett L, Egan LJ, Golden A, Ali RAR, Seoighe C. Gene-set analysis is severely biased when applied to genome-wide methylation data. *Bioinformatics* 2013, **29**(15): 1851-1857.
305. Trusch F, Klein M, Finsterbusch T, Kühn J, Hofmann J, Ehlers B. Seroprevalence of human polyomavirus 9 and cross-reactivity to African green monkey-derived lymphotropic polyomavirus. *Journal of General Virology* 2012, **93**(Pt 4): 698-705.
306. Tolstov YL, Pastrana DV, Feng H, Becker JC, Jenkins FJ, Moschos S, *et al.* Human Merkel cell polyomavirus infection II. MCV is a common human infection that can be detected by conformational capsid epitope immunoassays. *International journal of cancer* 2009, **125**(6): 1250-1256.
307. Knowles WA, Pipkin P, Andrews N, Vyse A, Minor P, Brown DW, *et al.* Population-based study of antibody to the human polyomaviruses BKV and JCV and the simian polyomavirus SV40. *Journal of medical virology* 2003, **71**(1): 115-123.
308. Robles C, Viscidi R, Malats N, Silverman DT, Tardon A, Garcia-Closas R, *et al.* Bladder cancer and seroreactivity to BK, JC and Merkel cell polyomaviruses: The Spanish bladder cancer study. *International Journal of Cancer* 2013, **133**(3): 597-603.
309. Gheit T, Muñoz JP, Levican J, González C, Ampuero S, Parra B, *et al.* Merkel cell polyomavirus in non-small cell lung carcinomas from Chile. *Experimental and molecular pathology* 2012, **93**(1): 162-166.
310. Xu S, Jiang J, Yu X, Sheng D, Zhu T, Jin M. Association of Merkel cell polyomavirus infection with EGFR mutation status in Chinese non-small cell lung cancer patients. *Lung Cancer* 2014, **83**(3): 341-346.
311. Arcavi L, Benowitz NL. Cigarette smoking and infection. *Archives of Internal Medicine* 2004, **164**(20): 2206-2216.
312. Van der Meijden E, Wunderink H, Van der Blij-de Brouwer C, Zaaijer H, Rotmans J, Bouwes Bavinck J. Human polyomavirus 9 infection in kidney transplant patients. *Emerging Infectious Diseases* 2014.
313. Hirsch HH, Knowles W, Dickenmann M, Passweg J, Klimkait T, Mihatsch MJ, *et al.* Prospective study of polyomavirus type BK replication and nephropathy in renal-transplant recipients. *New England Journal of Medicine* 2002, **347**(7): 488-496.
314. Hirsch H, Vincenti F, Friman S, Tuncer M, Citterio F, Wiecek A, *et al.* Polyomavirus BK replication in de novo kidney transplant patients receiving tacrolimus or cyclosporine: a prospective, randomized, multicenter study. *American Journal of Transplantation* 2013, **13**(1): 136-145.
315. Gardner S, MacKenzie E, Smith C, Porter A. Prospective study of the human polyomaviruses BK and JC and cytomegalovirus in renal transplant recipients. *Journal of clinical pathology* 1984, **37**(5): 578-586.

316. Antonsson A, Pawlita M, Feltkamp MC, Bouwes Bavinck JN, Euvrard S, Harwood CA*, et al.* Longitudinal study of seroprevalence and serostability of the human polyomaviruses JCV and BKV in organ transplant recipients. *Journal of medical virology* 2013, **85**(2)**:** 327-335.
317. Huang G, Chen LZ, Qiu J, Wang CX, Fei JG, Deng SX*, et al.* Prospective study of polyomavirus BK replication and nephropathy in renal transplant recipients in China: a single-center analysis of incidence, reduction in immunosuppression and clinical course. *Clinical transplantation* 2010, **24**(5)**:** 599-609.
318. Wiedinger K, Bitsaktsis C, Chang S. Reactivation of human polyomaviruses in immunocompromised states. *Journal of NeuroVirology* 2014**:** 1-8.
319. Costa C, Bergallo M, Ferrara P, Cavallo R. High-level polyomavirus JC viruria following long-term steroid therapy. *New Microbiologica* 2010, **33:** 405-407.
320. Girmanova E, Brabcova I, Klema J, Hribova P, Wohlfartova M, Skibova J*, et al.* Molecular networks involved in the immune control of BK polyomavirus. *Journal of Immunology Research* 2012, **2012**.
321. Neu U, Stehle T, Atwood WJ. The Polyomaviridae: Contributions of virus structure to our understanding of virus receptors and infectious entry. *Virology* 2009, **384**(2)**:** 389-399.
322. Pipas J. Common and unique features of T antigens encoded by the polyomavirus group. *Journal of virology* 1992, **66**(7)**:** 3979-3985.
323. Costa C, Saldan A, Cavallo R. Evaluation of virus-specific cellular immune response in transplant patients. *World journal of virology* 2012, **1**(6)**:** 150.
324. Costa C, Cavallo R. Polyomavirus-associated nephropathy. *World journal of transplantation* 2012, **2**(6)**:** 84.
325. Khanna N, Wolbers M, Mueller NJ, Garzoni C, Du Pasquier RA, Fux CA*, et al.* JC virus-specific immune responses in human immunodeficiency virus type 1 patients with progressive multifocal leukoencephalopathy. *Journal of Virology* 2009, **83**(9)**:** 4404-4411.
326. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M*, et al.* Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. *Virology* 2013, **436**(2)**:** 295-303.
327. Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U*, et al.* Identification of a novel human polyomavirus in organs of the gastrointestinal tract. *PloS One* 2013, **8**(3)**:** e58021.
328. Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K*, et al.* Identification of a novel polyomavirus in a pancreatic transplant recipient with retinal blindness and vasculitic myopathy. *Journal of Infectious Diseases* 2014**:** jiu250.
329. Accomando WP, Wiencke JK, Houseman EA, Nelson HH, Kelsey KT. Quantitative reconstruction of leukocyte subsets using DNA methylation. *Genome biology* 2014, **15**(3)**:** R50.
330. Pollack SV. Skin cancer in the elderly. *Clinics in geriatric medicine* 1987, **3**(4)**:** 715-728.
331. Kraemer KH, Lee M-M, Andrews AD, Lambert WC. The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer: the xeroderma pigmentosum paradigm. *Archives of dermatology* 1994, **130**(8)**:** 1018-1021.

332. Johnson TM, Rowe DE, Nelson BR, Swanson NA. Squamous cell carcinoma of the skin (excluding lip and oral mucosa). *Journal of the American Academy of Dermatology* 1992, **26**(3)**:** 467-484.
333. Moens U, Van Ghelue M, Ehlers B. Are human polyomaviruses co-factors for cancers induced by other oncoviruses? *Reviews in Medical Virology* 2014.
334. Rollison DE. Epidemiologic studies of polyomaviruses and cancer: previous findings, methodologic challenges and future directions. *Advances in experimental medicine and biology* 2006, **577:** 342-356.
335. International Agency for Research on Cancer. *Human papillomaviruses*, vol. 90. World Health Organization: Lyon, France, 2007.
336. Sweet BH, Hilleman MR. The vacuolating virus, S.V. 40. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine (New York, NY)* 1960, **105:** 420-427.
337. Poulin DL, DeCaprio JA. Is there a role for SV40 in human cancer? *Journal of clinical oncology* 2006, **24**(26)**:** 4356-4365.
338. Foulongne V, Kluger N, Dereure O, Brieu N, Guillot B, Segondy M. Merkel cell polyomavirus and Merkel cell carcinoma, France. *Emerging infectious diseases* 2008, **14**(9)**:** 1491.
339. Andersson K, Michael KM, Luostarinen T, Waterboer T, Gislefoss R, Hakulinen T*, et al.* Prospective study of human papillomavirus seropositivity and risk of nonmelanoma skin cancer. *American journal of epidemiology* 2012**:** kwr373.
340. Karagas MR, Tosteson TD, Blum J, Morris JS, Baron JA, Klaue B. Design of an epidemiologic study of drinking water arsenic exposure and skin and bladder cancer risk in a US population. *Environmental Health Perspectives* 1998, **106**(Suppl 4)**:** 1047–1050.
341. Gossai A, Waterboer T, Nelson HH, Michel A, Willhauck-Fleckenstein M, Farzan SF*, et al.* Seroepidemiology of Human Polyomaviruses in a US Population. *American journal of epidemiology* 2015.
342. Long EO, Kim HS, Liu D, Peterson ME, Rajagopalan S. Controlling NK Cell Responses: Integration of Signals for Activation and Inhibition. *Annual review of immunology* 2013, **31**.
343. Vineretsky KA, Karagas MR, Christensen BC, Kuriger-Laber JK, Perry AE, Storm CA*, et al.* Skin cancer risk is modified by KIR/HLA interactions that influence the activation of natural killer immune cells. *Cancer research* 2015 (in press).
344. Vineretsky KA, Karagas MR, Kuriger-Laber JK, Waterboer T, Pawlita M, Nelson HH. HLA-C-35kb Expression SNP Is Associated with Differential Control of $\beta$ -HPV Infection in Squamous Cell Carcinoma Cases and Controls. *PloS One* 2014, **9**(8)**:** e103710.
345. Vineretsky KA, Karagas MR, Pawlita M, Nelson HH. Functional killer-immunoglobulin-like receptor variation and risk of non-melanoma skin cancer. *Cancer Research* 2013, **73**(8 Supplement)**:** 1332-1332.
346. Ritchie MD, Hahn LW, Moore JH. Power of multifactor dimensionality reduction for detecting gene-gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity. *Genetic epidemiology* 2003, **24**(2)**:** 150-157.
347. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene–gene and gene–environment interactions. *Bioinformatics* 2003, **19**(3)**:** 376-382.

348. Moore JH, Gilbert JC, Tsai C-T, Chiang F-T, Holden T, Barney N, *et al.* A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *Journal of theoretical biology* 2006, **241**(2): 252-261.
349. Gui J, Andrew AS, Andrews P, Nelson HM, Kelsey KT, Karagas MR, *et al.* A simple and computationally efficient sampling approach to covariate adjustment for multifactor dimensionality reduction analysis of epistasis. *Human heredity* 2010, **70**(3): 219-225.
350. Csoma E, Mészáros B, Asztalos L, Kónya J, Gergely L. Prevalence of WU and KI polyomaviruses in plasma, urine, and respiratory samples from renal transplant patients. *Journal of medical virology* 2011, **83**(7): 1275-1278.
351. Kuypers J, Campbell AP, Guthrie KA, Wright NL, Englund JA, Corey L, *et al.* WU and KI polyomaviruses in respiratory samples from allogeneic hematopoietic cell transplant recipients. *Emerging infectious diseases* 2012, **18**(10).
352. Al Ustwani O, Kurzrock R, Wetzler M. Genetics on a WHIM. *British journal of haematology* 2014, **164**(1): 15-23.
353. Kawai T, Malech HL. WHIM syndrome: congenital immune deficiency disease. *Current opinion in hematology* 2009, **16**(1): 20-26.
354. Saundh BK, Tibble S, Baker R, Sasnauskas K, Harris M, Hale A. Different patterns of BK and JC polyomavirus reactivation following renal transplantation. *Journal of clinical pathology* 2010, **63**(8): 714-718.
355. Janeway C, Travers P, Walport M, Shlomchik M. Immunobiology: The Immune System in Health and Disease (Garland Science, New York). 2005.
356. Middleton D, Gonzelez F. The extensive polymorphism of KIR genes. *Immunology* 2010, **129**(1): 8-19.
357. Qi Y, Martin MP, Gao X, Jacobson L, Goedert JJ, Buchbinder S, *et al.* KIR/HLA pleiotropism: protection against both HIV and opportunistic infections. *PLoS Pathogens* 2006, **2**(8): e79.
358. Khakoo SI, Thio CL, Martin MP, Brooks CR, Gao X, Astemborski J, *et al.* HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science* 2004, **305**(5685): 872-874.
359. Purdy AK, Campbell KS. Natural killer cells and cancer: regulation by the killer cell Ig-like receptors (KIR). *Cancer Biology & Therapy* 2009, **8**(23): 2211-2220.
360. Blais ME, Dong T, Rowland-Jones S. HLA-C as a mediator of natural killer and T-cell activation: spectator or key player? *Immunology* 2011, **133**(1): 1-7.
361. Campillo JA, Martínez-Escribano JA, Moya-Quiles MR, Marín LA, Muro M, Guerra N, *et al.* Natural killer receptors on CD8 T cells and natural killer cells from different HLA-C phenotypes in melanoma patients. *Clinical cancer research* 2006, **12**(16): 4822-4831.
362. Hampras SS, Giuliano AR, Lin H-Y, Fisher KJ, Abrahamsen ME, McKay-Chopin S, *et al.* Natural history of polyomaviruses in men: The HIM Study. *Journal of Infectious Diseases* 2014: jiu626.
363. Wold WS, Mackey JK, Brackmann KH, Takemori N, Rigden P, Green M. Analysis of human tumors and human malignant cell lines for BK virus-specific DNA sequences. *Proceedings of the National Academy of Sciences* 1978, **75**(1): 454-458.
364. Mertz KD, Pfaltz M, Junt T, Schmid M, Figueras MTF, Pfaltz K, *et al.* Merkel cell polyomavirus is present in common warts and carcinoma in situ of the skin. *Human Pathology* 2010, **41**(10): 1369-1379.

365. Ganzenmueller T, Yakushko Y, Kluba J, Henke-Gendo C, Gutzmer R, Schulz TF. Next-generation sequencing fails to identify human virus sequences in cutaneous squamous cell carcinoma. *International Journal of Cancer* 2012, **131**(7): E1173-E1179.
366. Gellrich S, Schewe C, Sterry W, Lukowsky A. Absence of SV40 and other polyomavirus (JCV, BKV) DNA in primary cutaneous B cell lymphomas. *Journal of Investigative Dermatology* 2005, **124**(1): 278-279.
367. Moens U, Van Ghelue M, Ludvigsen M, Korup-Schulz S, Ehlers B. The early and late promoters of BKPyV, MCPyV, TSPyV, and HPyV12 are among the strongest of all known human polyomaviruses in 10 different cell lines. *The Journal of General Virology* 2015.
368. Borchert S, Czech-Sioli M, Neumann F, Schmidt C, Wimmer P, Dobner T, *et al.* High-affinity Rb binding, p53 inhibition, subcellular localization, and transformation by wild-type or tumor-derived shortened Merkel cell polyomavirus large T antigens. *Journal of virology* 2014, **88**(6): 3144-3160.
369. Sullivan CS, Grundhoff AT, Tevethia S, Pipas JM, Ganem D. SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* 2005, **435**(7042): 682-686.
370. Gordon J, Khalili K. The human polyomavirus, JCV, and neurological diseases (review). *International journal of molecular medicine* 1998, **1**(4): 647-702.
371. Stettner MR, Nance JA, Wright CA, Kinoshita Y, Kim W-K, Morgello S, *et al.* SMAD proteins of oligodendroglial cells regulate transcription of JC virus early and late genes coordinately with the Tat protein of human immunodeficiency virus type 1. *Journal of General Virology* 2009, **90**(8): 2005-2014.
372. Atencio I, Belli B, Hobbs M, Cheng S, Villarreal L, Fan H. A model for mixed virus disease: co-infection with Moloney murine leukemia virus potentiates runting induced by polyomavirus (A2 strain) in Balb/c and NIH Swiss mice. *Virology* 1995, **212**(2): 356-366.
373. Antonsson A, Green AC, Mallitt KA, O'Rourke PK, Pawlita M, Waterboer T, *et al.* Prevalence and stability of antibodies to the BK and JC polyomaviruses: a long-term longitudinal study of Australians. *The Journal of General Virology* 2010, **91**(Pt 7): 1849-1853.
374. Bressollette-Bodin C, Coste-Burel M, Hourmant M, Sebille V, Andre-Garnier E, Imbert-Marcille BM. A prospective longitudinal study of BK virus infection in 104 renal transplant recipients. *American Journal of Transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 2005, **5**(8): 1926-1933.
375. Doucette KE, Pang XL, Jackson K, Burton I, Carbonneau M, Cockfield S, *et al.* Prospective monitoring of BK polyomavirus infection early posttransplantation in nonrenal solid organ transplant recipients. *Transplantation* 2008, **85**(12): 1733-1736.
376. Razonable RR, Brown RA, Humar A, Covington E, Alecock E, Paya CV. A longitudinal molecular surveillance study of human polyomavirus viremia in heart, kidney, liver, and pancreas transplant patients. *The Journal of Infectious Diseases* 2005, **192**(8): 1349-1354.
377. Jilek S, Jaquiery E, Hirsch HH, Lysandropoulos A, Canales M, Guignard L, *et al.* Immune responses to JC virus in patients with multiple sclerosis treated with natalizumab: a cross-sectional and longitudinal study. *The Lancet Neurology* 2010, **9**(3): 264-272.

378. Trampe AK, Hemmelmann C, Stroet A, Haghikia A, Hellwig K, Wiendl H, *et al.* Anti-JC virus antibodies in a large German natalizumab-treated multiple sclerosis cohort. *Neurology* 2012, **78**(22)**:** 1736-1742.
379. Ryschkewitsch CF, Jensen PN, Monaco MC, Major EO. JC virus persistence following progressive multifocal leukoencephalopathy in multiple sclerosis patients treated with natalizumab. *Annals of Neurology* 2010, **68**(3)**:** 384-391.
380. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence Estimate of Nonmelanoma Skin Cancer (Keratinocyte Carcinomas) in the US Population, 2012. *JAMA Dermatology* 2015.
381. Rogers HW, Weinstock MA, Harris AR, Hinckley MR, Feldman SR, Fleischer AB, *et al.* Incidence estimate of nonmelanoma skin cancer in the United States, 2006. *Archives of Dermatology* 2010, **146**(3)**:** 283-287.
382. Greenberg ER, Baron JA, Stevens MM, Stukel TA, Mandel JS, Spencer SK, *et al.* The Skin Cancer Prevention Study: design of a clinical trial of beta-carotene among persons at high risk for nonmelanoma skin cancer. *Controlled clinical trials* 1989, **10**(2)**:** 153-166.
383. Karagas MR, Greenberg ER, Nierenberg D, Stukel TA, Morris JS, Stevens MM, *et al.* Risk of squamous cell carcinoma of the skin in relation to plasma selenium, alpha-tocopherol, beta-carotene, and retinol: a nested case-control study. *Cancer Epidemiology Biomarkers & Prevention* 1997, **6**(1)**:** 25-29.
384. Fleiss J. The Design and Analysis of Clinical Experiments. Taylor & Francis; 1986.
385. Antonsson A, Green AC, Mallitt K-a, O'Rourke PK, Pandeya N, Pawlita M, *et al.* Prevalence and stability of antibodies to 37 human papillomavirus types—a population-based longitudinal study. *Virology* 2010, **407**(1)**:** 26-32.
386. Fleiss JL, Levin BR, Paik MC. Statistical methods for rates and proportions. *Hoboken, NJ: J Wiley-Interscience* 2003, **203:** 151.
387. Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960, **20**(1)**:** 37-46.
388. Rothman K. *Epidemiology An Introduction*. Oxford University Press: London, 2002.
389. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2011, **73**(1)**:** 3-36.
390. Wood S. *Generalized additive models: an introduction with R*. CRC press, 2006.
391. Randhawa P, Bohl D, Brennan D, Ruppert K, Ramaswami B, Storch G, *et al.* Longitudinal analysis of levels of immunoglobulins against BK virus capsid proteins in kidney transplant recipients. *Clinical and Vaccine Immunology : CVI* 2008, **15**(10)**:** 1564-1571.
392. Iannetta M, Bellizzi A, Lo Menzo S, Anzivino E, D'Abramo A, Oliva A, *et al.* HIV-associated progressive multifocal leukoencephalopathy: longitudinal study of JC virus non-coding control region rearrangements and host immunity. *Journal of Neurovirology* 2013, **19**(3)**:** 274-279.
393. Flo RW, Nilsen A, Voltersvik P, Haukenes G. Serum antibodies to viral pathogens and Toxoplasma gondii in HIV-infected individuals. *Acta Pathologica, Microbiologica et Immunologica Scandinavica* 1993, **101**(12)**:** 946-952.
394. Plavina T, Subramanyam M, Bloomgren G, Richman S, Pace A, Lee S, *et al.* Anti-JC virus antibody levels in serum or plasma further define risk of natalizumab-associated progressive multifocal leukoencephalopathy. *Annals of Neurology* 2014, **76**(6)**:** 802-812.

395. Dominguez-Mozo MI, Garcia-Montojo M, De Las Heras V, Garcia-Martinez A, Arias-Leal AM, Casanova I, *et al.* Anti-JCV antibodies detection and JCV DNA levels in PBMC, serum and urine in a cohort of Spanish Multiple Sclerosis patients treated with natalizumab. *Journal of Neuroimmune Pharmacology : the official journal of the Society on NeuroImmune Pharmacology* 2013, **8**(5): 1277-1286.
396. Lin J, Bettin P, Lee JK, Ho JK, Sadiq SA. Cerebrospinal fluid and serum JC virus antibody detection in multiple sclerosis patients treated with natalizumab. *Journal of Neuroimmunology* 2013, **261**(1-2): 123-128.
397. Rollison DE, Engels EA, Halsey NA, Shah KV, Viscidi RP, Helzlsouer KJ. Prediagnostic circulating antibodies to JC and BK human polyomaviruses and risk of non-Hodgkin lymphoma. *Cancer Epidemiology Biomarkers & Prevention* 2006, **15**(3): 543-550.
398. Bouvard V, Baan RA, Grosse Y, Lauby-Secretan B, El Ghissassi F, Benbrahim-Tallaa L, *et al.* Carcinogenicity of malaria and of some polyomaviruses. *The Lancet Oncology* 2012, **13**(4): 339-340.
399. Lassak A, Del Valle L, Peruzzi F, Wang JY, Enam S, Croul S, *et al.* Insulin receptor substrate 1 translocation to the nucleus by the human JC virus T-antigen. *Journal of Biological Chemistry* 2002, **277**(19): 17231-17238.
400. Enam S, Del Valle L, Lara C, Gan D-D, Ortiz-Hidalgo C, Palazzo JP, *et al.* Association of Human Polyomavirus JCV with Colon Cancer Evidence for Interaction of Viral T-Antigen and β-Catenin. *Cancer Research* 2002, **62**(23): 7093-7101.
401. Gan D-D, Reiss K, Carrill T, Del Valle L, Croul S, Giordano A, *et al.* Involvement of Wnt signaling pathway in murine medulloblastoma induced by human neurotropic JC virus. *Oncogene* 2001, **20**(35): 4864-4870.
402. Trojanek J, Ho T, Del Valle L, Nowicki M, Wang JY, Lassak A, *et al.* Role of the insulin-like growth factor I/insulin receptor substrate 1 axis in Rad51 trafficking and DNA repair by homologous recombination. *Molecular and Cellular Biology* 2003, **23**(21): 7510-7524.
403. Michaud DS, Houseman EA, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. Understanding the role of the immune system in the development of cancer: new opportunities for population-based research. *Cancer Epidemiology Biomarkers & Prevention* 2015, **24**(12): 1811-1819.
404. Casabonne D, Michael KM, Waterboer T, Pawlita M, Forslund O, Burk RD, *et al.* A prospective pilot study of antibodies against human papillomaviruses and cutaneous squamous cell carcinoma nested in the Oxford component of the European Prospective Investigation into Cancer and Nutrition. *International Journal of Cancer* 2007, **121**(8): 1862-1868.
405. Engels EA, Wacholder S, Katki HA, Chaturvedi AK. Tumor-Based Case–Control Studies of Infection and Cancer: Muddling the When and Where of Molecular Epidemiology. *Cancer Epidemiology Biomarkers & Prevention* 2014, **23**(10): 1959-1964.
406. Wijburg MT, Siepman D, van Eijk JJ, Killestein J, Wattjes MP. Concomitant granule cell neuronopathy in patients with natalizumab-associated PML. *Journal of Neurology* 2016: 1-8.