SUPPLEMENTARY DATA

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

List of Supplemental Figures:

Supplemental Figure 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 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 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. 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 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

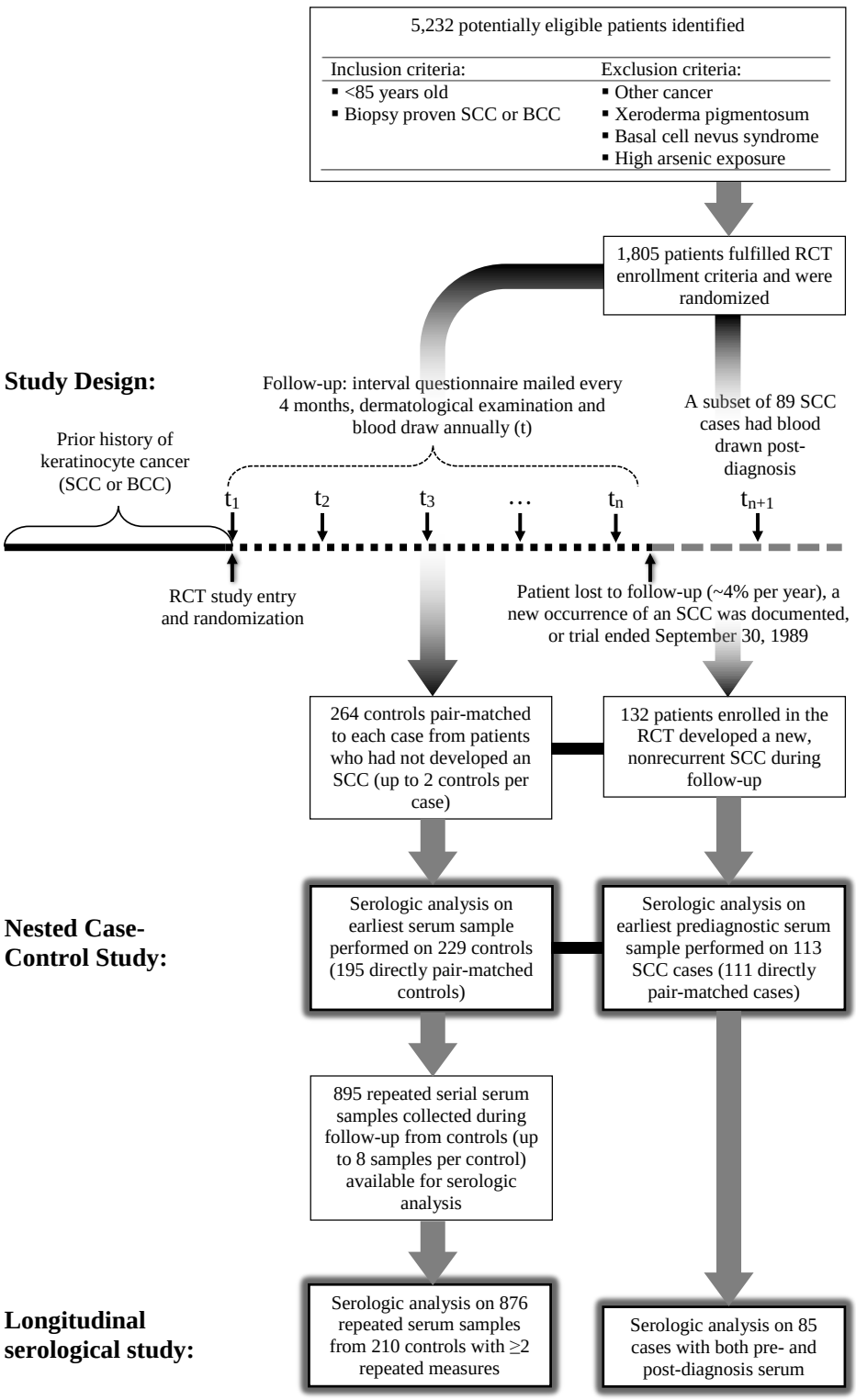
a prior SCC (n=186 controls and 105 cases for conditional analyses; n=220 controls and 107 cases for unconditional analyses)), and **C.** having ever had a prior BCC ("BCC" refers to having had a prior BCC (n=100 controls and 68 cases for conditional analyses; n=118 controls and 69 cases for unconditional 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 and Supplemental Table 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 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 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.

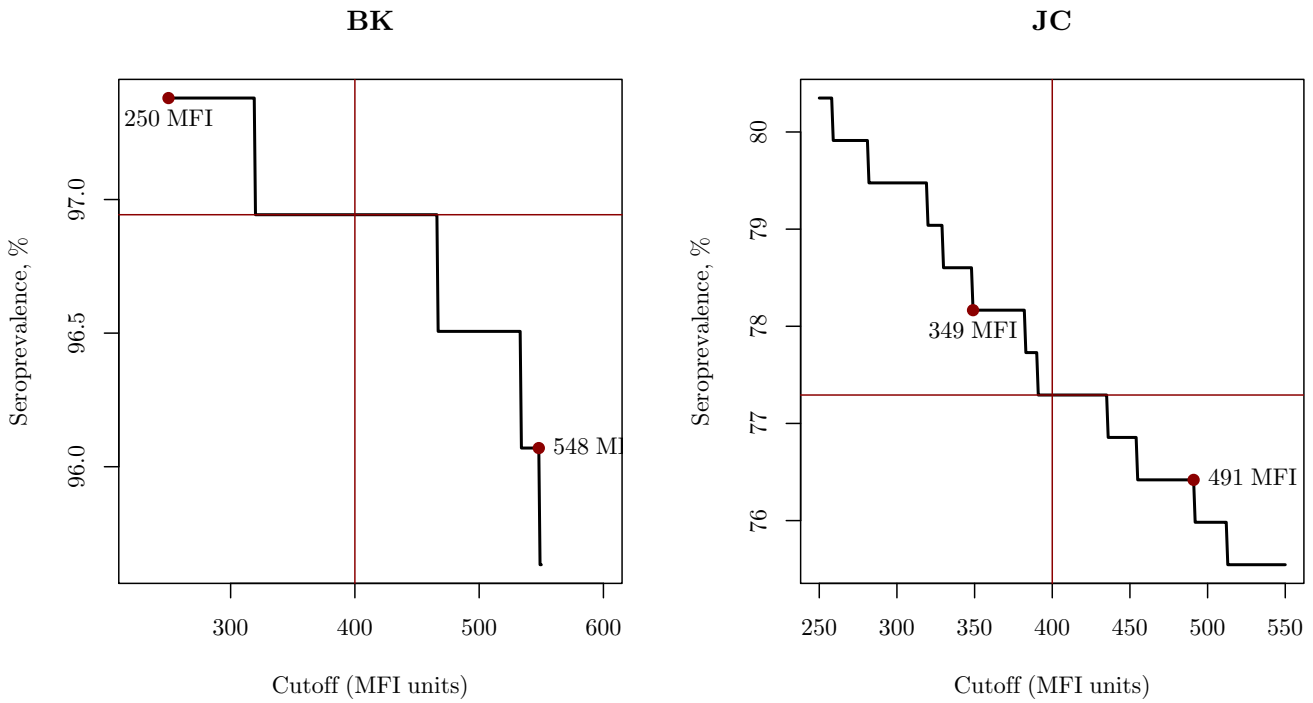
List of Supplemental Tables:

Supplemental Table 1. Distribution of BK and JC human polyomavirus (PyV) seropositivity by selected baseline characteristics among 229 controls from the Skin Cancer Prevention Study.

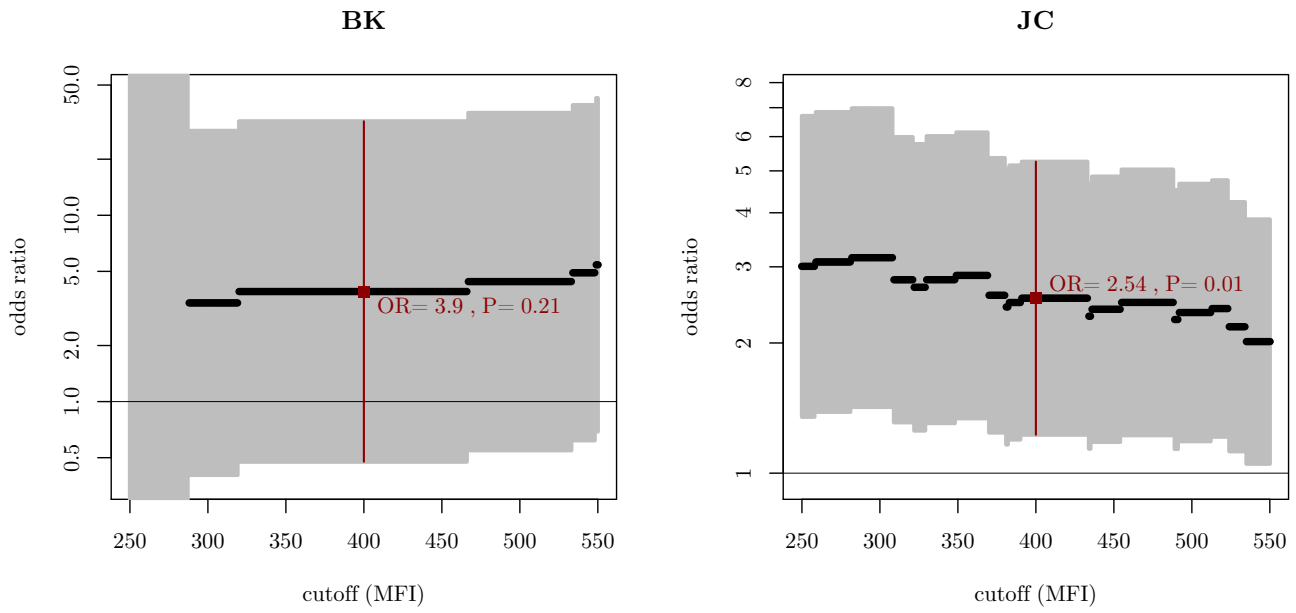
Supplemental Table 2. Unconditional odds ratios (95% confidence intervals) for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type and quartiles of PyV seroreactivity at baseline among 342 study participants from the Skin Cancer Prevention Study.



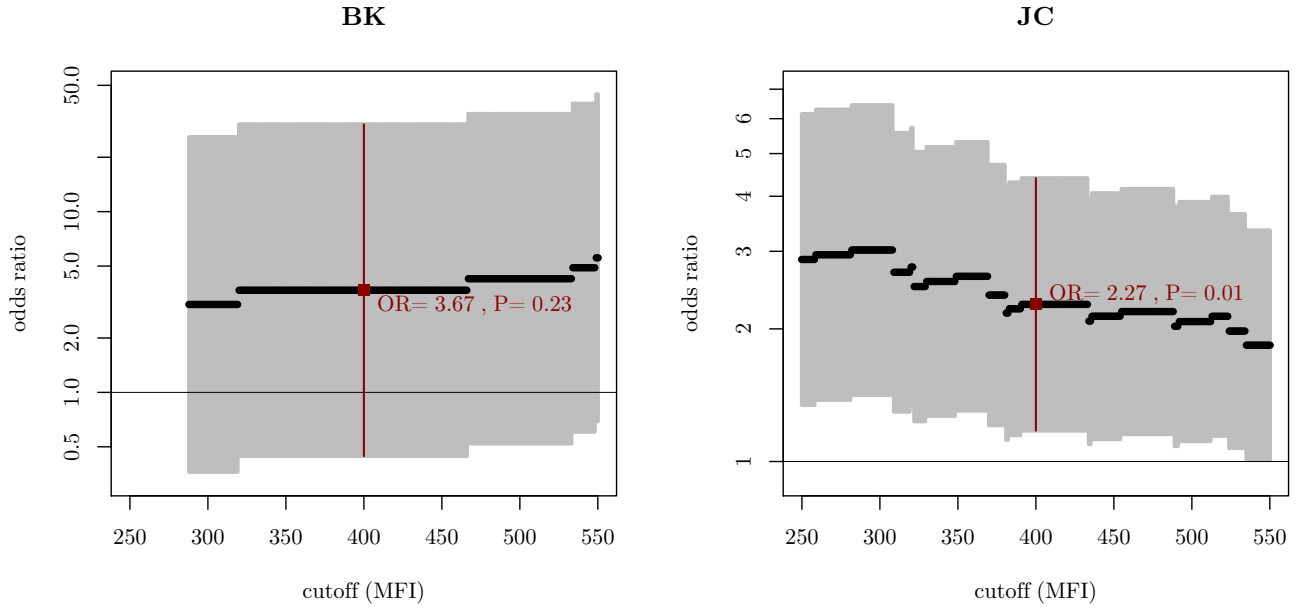
Supplemental Figure 1. Flowchart of the "Skin Cancer Prevention Study" randomized clinical trial (RCT) design from which the study group was derived. Patients with a prior history of squamous cell carcinoma (SCC) or basal cell carcinoma (BCC) were enrolled. In a nested case-control study, risk of a new, 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.



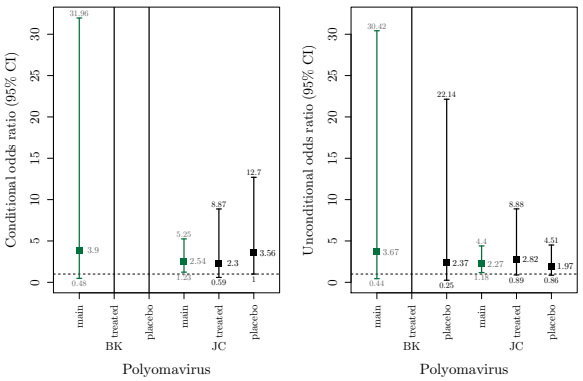
Supplemental Figure 2. Robustness of human polyomavirus (PyV) seroprevalence estimates among 229 controls from the Skin Cancer Prevention Study. The cut points were varied from 250 to 550 median fluorescence intensity (MFI) units (x axis), and the resulting seroprevalences were calculated using the new 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.



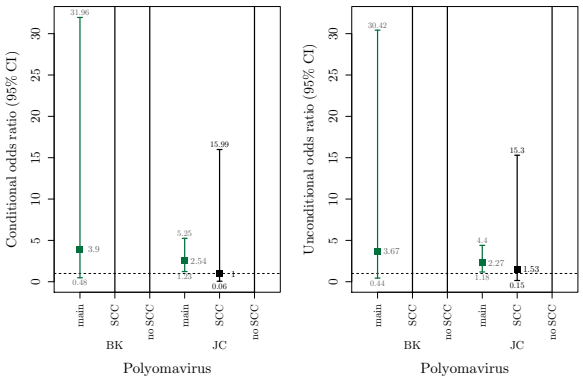
Supplemental Figure 3. Robustness of odds ratio (OR) estimates obtained from conditional logistic regression analysis for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among 306 study participants from the Skin Cancer Prevention Study. The cut points were 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.



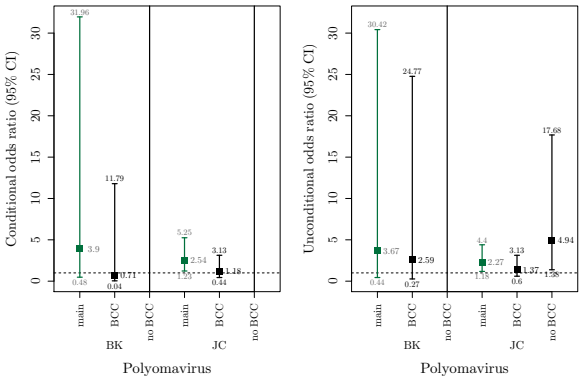
Supplemental Figure 4. Robustness of odds ratio (OR) estimates obtained from unconditional logistic regression analysis for cutaneous squamous cell carcinoma (SCC) by seropositivity for each polyomavirus (PyV) type among 342 study participants from the Skin Cancer Prevention Study, with adjustment for 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.



Sub-Figure A

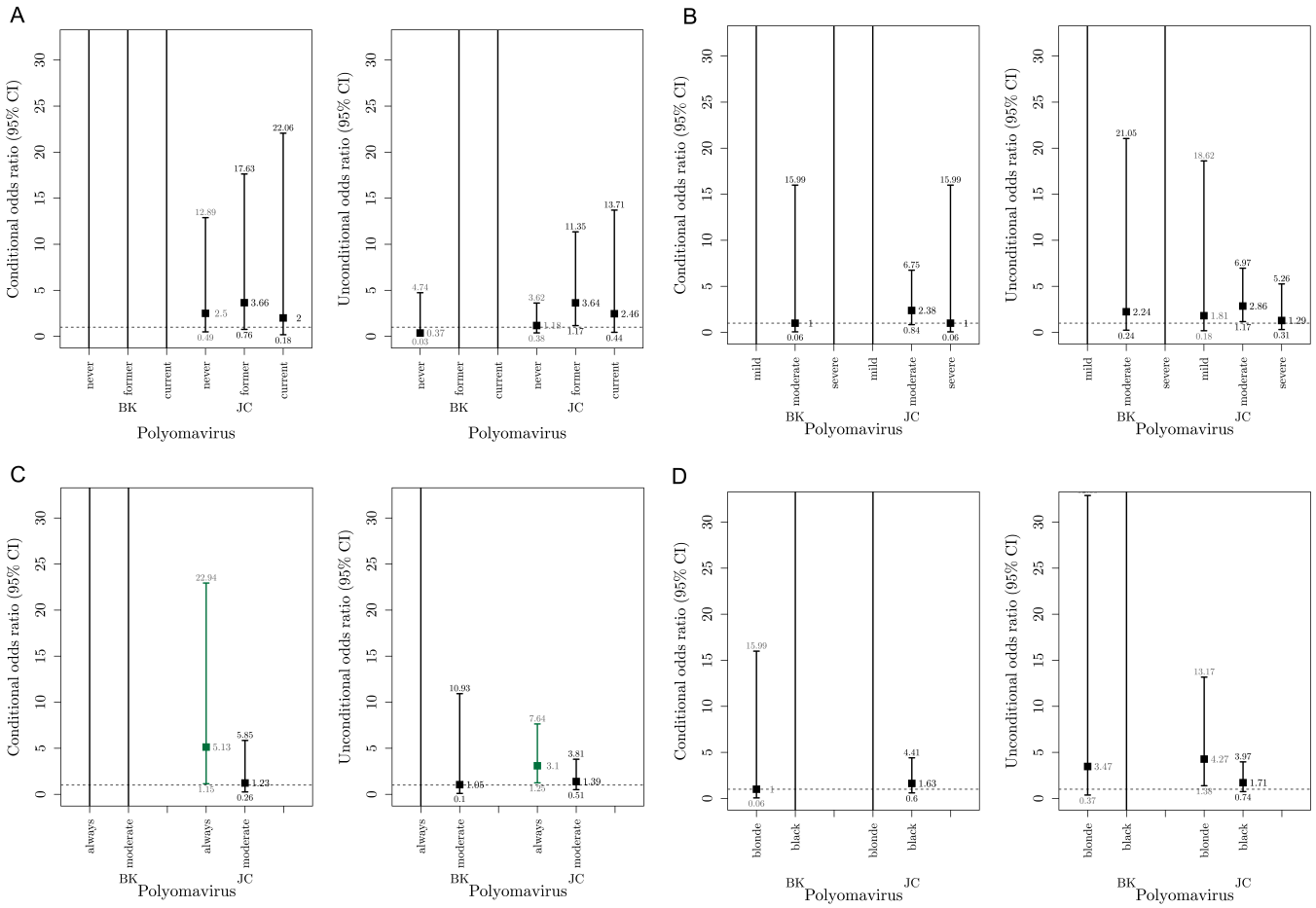


Sub-Figure B



Sub-Figure C

Supplemental Figure 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 and Supplemental Table 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 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 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.

Supplemental Table 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 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.