

**Keywords:** height; prospective cohort study; basal cell carcinoma; squamous cell carcinoma; height-related SNPs; genetic score

# Height, height-related SNPs, and risk of non-melanoma skin cancer

Xin Li<sup>1</sup>, Liming Liang<sup>1</sup>, Yen-Chen Anne Feng<sup>1</sup>, Immaculata De Vivo<sup>1,2</sup>, Edward Giovannucci<sup>1,2,3</sup>, Jean Y Tang<sup>4</sup> and Jiali Han<sup>\*,5,6,7</sup>

<sup>1</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA; <sup>2</sup>Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA; <sup>3</sup>Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA; <sup>4</sup>Department of Dermatology, Stanford University School of Medicine, Redwood City, CA, USA; <sup>5</sup>Department of Epidemiology, Fairbanks School of Public Health, Indiana University, Indianapolis, IN, USA; <sup>6</sup>Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, IN, USA and <sup>7</sup>Center for Pharmacoepidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN, USA

**Background:** Adult height has been associated with risk of several site-specific cancers, including melanoma. However, less attention has been given to non-melanoma skin cancer (NMSC).

**Methods:** We prospectively examined the risk of squamous cell carcinoma (SCC) and basal cell carcinoma (BCC) in relation to adult height in the Nurses' Health Study (NHS,  $n = 117\,863$ ) and the Health Professionals Follow-up Study (HPFS,  $n = 51\,111$ ). We also investigated the relationships between height-related genetic markers and risk of BCC and SCC in the genetic data sets of the NHS and HPFS (3898 BCC cases, and 8530 BCC controls; 527 SCC cases, and 8962 SCC controls).

**Results:** After controlling for potential confounding factors, the hazard ratios were 1.09 (95% CI: 1.02, 1.15) and 1.10 (95% CI: 1.07, 1.13) for the associations between every 10 cm increase in height and risk of SCC and BCC respectively. None of the 687 height-related single-nucleotide polymorphisms (SNPs) was significantly associated with the risk of SCC or BCC, nor were the genetic scores combining independent height-related loci.

**Conclusions:** Our data from two large cohorts provide further evidence that height is associated with an increased risk of NMSC. More studies on height-related genetic loci and early-life exposures may help clarify the underlying mechanisms.

Non-melanoma skin cancer (NMSC), including basal and squamous cell carcinomas (BCC and SCC, respectively), is the most common malignancy among Caucasians (Narayanan *et al*, 2010). It is estimated that over 2 million cases of NMSC occur each year in the US, with incidence continuing to increase (Rogers *et al*, 2010). BCC rarely metastasises to other organs or causes death; however, this malignancy results in considerable morbidity and places a huge burden on the health-care system worldwide (Tung and Vidimos, 2002). In contrast, SCC is more likely to invade other tissues and can be fatal (Tung and Vidimos, 2002). Both environmental and constitutional factors contribute to the development of NMSC. Ultraviolet radiation is a well-established carcinogen for both BCC and SCC (Gallagher *et al*, 1995a;

Armstrong *et al*, 1997). Constitutional risk factors that represent certain components of genetic susceptibility include hair colour, family history, tanning ability, and so forth (Gallagher *et al*, 1995a, b; Han *et al*, 2006).

Taller people are more likely to develop cancer (Renehan, 2011). Though a number of case-control (Cutler *et al*, 1996; Shors *et al*, 2001; Gallus *et al*, 2006; Olsen *et al*, 2008) and cohort studies (Thune *et al*, 1993; Green *et al*, 2011; Kabat *et al*, 2013a, b; Wirén *et al*, 2014) have examined the association between adult height and risk of melanoma skin cancer, the association between height and risk of NMSC has been infrequently investigated. Though one prospective study reported a significantly higher risk of NMSC among taller men and women (Wirén *et al*, 2014), BCC and SCC

\*Correspondence: Dr J Han; E-mail: jialhan@iu.edu

Received 1 July 2016; revised 4 October 2016; accepted 12 October 2016; published online 15 November 2016

© 2017 Cancer Research UK. All rights reserved 0007–0920/17

were not analysed separately. In addition, the authors failed to consider important confounders such as race, constitutional factors, and sun exposure history, and potential effect modifications by them. Therefore, a comprehensive assessment of the relationships between height and risk of different types of NMSC is still lacking.

The underlying mechanism for this positive association remains unclear. One possible explanation is that height-related genetic factors are also tied to skin cancer; however, studies exploring this possibility are rare. Adult height is determined by genetic factors to a great extent (Yang *et al*, 2010). The largest genome-wide association study (GWAS,  $n = 253\,288$ ) on height was conducted by the Genetic Investigation of Anthropometric Traits (GIANT) consortium, which identified 697 variants at genome-wide significance that together explain one-fifth of the heritability for adult height (Wood *et al*, 2014). Testing the associations between these height-related single-nucleotide polymorphisms (SNPs) and NMSC risk may help us better understand the relationship between these two phenotypes and provide more insight into skin tumorigenesis.

Here we used data from the Nurses' Health Study (NHS) and the Health Professionals Follow-up Study (HPFS) to investigate the association between height and risk of incident SCC and BCC simultaneously. We also evaluated the extent to which the observed associations were affected by confounding factors, and tested potential interactions between height and other factors on NMSC risk. To better understand the association at the genetic level, we also examined the individual and combined associations of height-related variants identified by the GIANT consortium with risk of NMSC in the genetic data sets of the NHS and HPFS.

## MATERIALS AND METHODS

### Study population

**Nurses' health study (NHS).** The NHS is a prospective cohort study established in 1976 with 121 700 female US registered nurses, who were then 30–55 years old. All of them completed and returned a mailed self-administered questionnaire about their medical histories and baseline lifestyle. In 1989 and 1990, a total of 32 826 women provided blood samples. Information regarding medical history, lifestyle, and disease diagnoses was updated every 2 years with a follow-up rate of 90%.

**Health professionals follow-up study (HPFS).** The HPFS began in 1986 with 51 529 US male health professionals who were 40–75 years old at initial recruitment. They all answered a detailed mailed questionnaire at the inception of the study. Disease- and health-related information was obtained and updated through biennial questionnaires. Between 1993 and 1994, 18 159 of these men provided a blood sample. The average follow-up rate for this cohort over 10 years is greater than 90%.

**Genetic data sets.** Eighteen case-control studies nested within the NHS and HPFS with cleaned genotype data were included. Samples from the 18 studies were genotyped using a variety of platforms, which we then combined into three compiled data sets based on their genotype platform types: Affymetrix (Affy), Illumina HumanHap series (Illumina), or Illumina Omni Express (Omni) (Supplementary Table 1). Quality control on SNP completion rate, sample completion rate, ancestry consistency, deviation from Hardy-Weinberg equilibrium, Mendelian consistency, minor allele frequency, and duplication were conducted within each of the three combined data sets. We then imputed the compiled data sets using the 1000 Genomes Project ALL Phase I Integrated Release Version 3 Haplotypes excluding monomorphic

and singleton sites (2010–11 data freeze, 2012-03-14 haplotypes) as the reference panel (Supplementary Table 2). Basic information on the 18 studies and detailed descriptions of quality control and imputation are provided in Supplementary Materials.

**Measurement of height and ascertainment of skin cancers.** Height was reported by participants at recruitment (1976 for NHS, 1986 for HPFS). New diagnoses of NMSC were reported by participants biennially. With their permission, participants' medical records were obtained and reviewed by physicians to confirm diagnoses of SCC. Medical records were not obtained for BCC; however, several studies support the validity of self-report of BCC in our cohorts. Colditz *et al* (1986) evaluated the validity of self-reported illnesses including skin cancer in the NHS. Among 33 random samples of women who had reported NMSC, medical records indicated that 30 (91%) had correctly reported their skin cancer. Also, Hunter *et al* (1990) previously examined risk factors for BCC in the NHS using self-reported cases. As expected, they found that lighter pigmentation and higher tendency to sunburn were associated with an increased risk of BCC. In addition, using the self-reported BCC cases, our group identified the previously well-documented genetic variant in the MC1R gene as the top risk locus in our GWAS for BCC (Nan *et al*, 2011).

**Measurement of covariates.** Information on skin cancer risk factors was obtained from questionnaires in both the NHS and the HPFS in the 1980s. The risk factors included: (1) natural hair colour at age 20; (2) family history of melanoma in first-degree relatives; (3) skin reaction after 2 h of sun exposure as a child/adolescent; (4) number of severe sunburns over lifetime; (5) mole count measuring 3 mm or larger on the left arm; and (6) states lived in at birth, age 15, and age 30.

Data on weight, smoking status, and menopausal status was first collected at baseline (1976 for NHS and 1986 for HPFS) and then updated biennially in subsequent questionnaires for all cohort members. Body mass index (BMI) was computed as weight in kilograms divided by the square of height in meters for each follow-up cycle. Physical activity was first asked with detail in 1986 in both cohorts and updated every 2 years thereafter. The reproducibility and validity of self-reported physical activity in both cohorts has been evaluated in detail in previous studies (Wolf *et al*, 1994; Chasan-Taber *et al*, 1996). Energy expenditure in metabolic equivalent tasks (METs) (Ainsworth *et al*, 1993) measured in hours per week was calculated by multiplying the number of hours per week of leisure-time physical activity by the metabolic equivalent (MET) value of the activity and summing the products of all types of activities. Food frequency questionnaires were initially collected in 1980 for the NHS and 1986 for the HPFS, and alcohol intake and diet were generally updated every 4 years. Previous studies have shown that the food-frequency questionnaire validly assesses dietary and alcohol intake during the past year (Willett *et al*, 1988; Salvini *et al*, 1989). Self-reported race that was measured in 1982 in NHS and 1986 in HPFS was also considered. Non-whites were collapsed into one group because of insufficient sample sizes in individual race categories.

**Height-related SNPs and calculation of genetic score.** Of 697 height-related SNPs identified by the GIANT consortium, 687 were available in our genetic data set. For a locus in which multiple SNPs in linkage disequilibrium (LD, defined as  $r^2 > 0.1$ ) were identified, we selected the SNP with the most significant association with height as reported by the GIANT paper, yielding 593 SNPs for genetic score calculation. The scores were calculated only for individuals who had no missing value in any of the chosen SNPs. We assumed an additive genetic model for each SNP, which performs well even when the true genetic model is unknown or wrongly specified (Balding, 2006). For each individual, we summed the dosage of alleles that are related to increase in height of the 593

SNPs to obtain the simple count genetic score. We also constructed a weighted score by multiplying the dosage of effect alleles by the corresponding regression coefficients in the original GWAS paper and then summing the products. Both the original simple count score and the weighted score were rescaled to a mean of 1186 alleles (2 alleles  $\times$  593 SNPs) before testing their associations with NMSC to make the results comparable. We presented the formula for calculating genetic scores in Supplementary Table 3.

### Statistical analysis

**Height and skin cancer.** Participants who did not report their date of birth or height were excluded, as were those who had invalid information on height at recruitment (i.e., whose reported height was  $<120$  or  $>200$  cm). Participants who had baseline cancers were excluded, and those who reported any type of cancer or died during follow-up were censored. We used Cox proportional hazards models stratified by follow-up cycles and age to calculate the hazard ratios (HRs) and 95% confidence intervals (CIs) of each type of skin cancer. Person-time was calculated for each participant from the date of baseline questionnaire return to the date of the first report of NMSC, death, or the end of follow-up (June 2010), whichever came first. When quantifying the relationship between NMSC and height, we modelled height as a continuous measure expressed in 10 cm (increasing) increments. In the multivariate analysis, we simultaneously controlled for age, smoking status, alcohol intake, BMI, physical activity, and menopausal status/postmenopausal hormone use (only in NHS). Then we fitted a more complex model by additionally including hair colour, family history of melanoma, sunburn reaction as a child/adolescent, number of severe sunburns, mole count, and states lived in at birth, age 15, and age 30. Lastly, race was controlled for in the model to assess potential confounding effects. We tested the heterogeneity of the results among men and women and conducted a fixed-effect meta-analysis if there was no significant gender difference. Multiplicative interactions between height and other potential risk factors of NMSC were tested using the likelihood ratio test comparing a 'main effect only' model vs a model with the product term. All covariates in the multivariable-adjusted models were considered and sequentially tested for interaction each at a time. All statistical analyses were performed using SAS software (version 9.3 for UNIX; SAS Institute, Cary, NC, USA). We considered two-sided *P*-values less than 0.05 to be statistically significant.

**Height-related SNPs and skin cancer.** Data on participants who appeared in more than one of the three combined data sets were included only once in analyses. Baseline common cancer cases were excluded, as were NMSC cases who had other common cancers before diagnosis of skin cancer. Eligible controls were free of skin cancers or other common cancers. We assessed the associations between individual height-related SNPs and SCC as well as BCC using logistic regression models adjusted for gender, age, and the top three eigenvectors. The same models were fitted for the associations between genetic scores and risk of NMSC. All the analyses were first conducted within each of the platform-specific data sets, and then combined by meta-analysis if results were not significantly different. ProbABEL package and R-3.0.2 were used to perform these tests. We considered two-sided *P*-values less than 0.05 to be statistically significant. Bonferroni correction using the number of independent tests was applied to account for multiple comparisons.

**Sensitivity analysis and validation of self-reported ancestry.** Ancestry within the white population is a potential confounder that may bias the estimation of height-skin cancer association. Height varies across Europe, with Northern Europeans generally taller than Southern Europeans (Cavelaars *et al*, 2000; Turchin *et al*, 2012; Gragstruber *et al*, 2014). Intra-European ethnic origin

has also been found to be related to both melanoma and NMSCs (D'Arcy *et al*, 1984; English *et al*, 1998). We adjusted self-reported race (Southern European/Mediterranean; Scandinavian; Other Caucasian; and Non-white ancestry) in the multivariable models for height-NMSC association; however, such information may be inaccurate.

Therefore, we used participants' genetic data to estimate their accurate ancestry. Genetic ancestry was represented by ancestry coordinates calculated by the Locating Ancestry from Sequence Reads (LASER) method, which has been demonstrated to accurately infer worldwide continental ancestry and even fine-scale ancestry within Europe. Detailed descriptions of LASER have been published previously (Wang *et al*, 2014, 2015). We tested the correlation between self-reported European ancestry and the first as well as the second ancestry coordinates to validate the information collected by the questionnaire. We also conducted a sensitivity analysis in which we compared the Cox models without ancestry, with self-reported ancestry, and with genetic ancestry coordinates as covariates. These analyses were restricted to participants in the genetic data set, all of whom are of European ancestry.

## RESULTS

**Height and skin cancer risk.** We included 117 863 and 51 111 participants from the NHS and the HPFS, respectively. We documented 1632 SCCs over 3 194 911 person-years and 21 366 BCCs during 3 183 210 person-years in the NHS. In the HPFS, 1259 SCC events during 869 263 person-years and 9715 BCCs over 860 910 person-years of follow-up were identified.

The baseline age-standardised characteristics of participants by quartiles of height are listed in Table 1. Taller participants tended to be younger, drank more alcohol, excised more, and were more likely to be current smokers. Higher prevalence of Scandinavian ethnicity, family history of melanoma, red/blond hair, presence of arm moles, and painful burn/blister skin reaction after prolonged sun exposure as a child/adolescent were found in higher quartiles of height. Study participants with short stature had a higher BMI than taller participants. These trends were consistent in men and women. In the NHS, the percentage of current hormone replacement therapy users is higher among taller women.

In the age-adjusted models (Model 1 in Table 2) and multivariate models without race (Models 2 and 3), height was significantly positively associated with risk of SCC and BCC in both men and women. Further including self-reported race (Model 4) did not alter the results materially in the NHS. Risk of SCC showed only a borderline association with height in the HPFS. In the full model (Model 4), HRs for the associations between per 10 cm increase in height and SCC were 1.09 (95% CI: 1.00, 1.18) in both women and men. For BCC, the HRs were 1.11 (95% CI: 1.09, 1.14) and 1.08 (95% CI: 1.05, 1.12), respectively, among females and males. Although the magnitude of association for BCC appeared to be slightly higher in the NHS, heterogeneity between genders did not reach statistical significance (Table 2). The combined HRs were 1.09 (95% CI: 1.02, 1.15) and 1.10 (95% CI: 1.07, 1.13) for the associations of each 10 cm increase in height with risk of SCC and BCC, respectively. We found no significant interaction between height and other covariates in the full multivariable-adjusted model.

**Height-related SNPs and skin cancer risk.** Sample sizes of the genetic data sets before exclusion and number of NMSC cases and controls after exclusion are shown in Table 3. Among the 687 height-related SNPs available in our genetic data set, 37 and 38 showed nominally significant associations (*P*-value  $<0.05$ ) with

**Table 1. Baseline characteristics by quartiles of height in the NHS (1976–2010) and HPFS (1986–2010)**

	Quartiles of height in cm							
	NHS (Female)				HPFS (Male)			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Age (year) <sup>a</sup> , mean (s.d.)	42.8(7.3)	42.5(7.3)	42.2(7.2)	41.7(7.1)	56.3(10.2)	54.7(9.9)	53.8(9.6)	52.4(9.3)
Height (cm), mean (s.d.)	155.6 (2.5)	161.5 (1.3)	166.4 (1.3)	172.5 (3.0)	169.4 (5.5)	176.8 (1.3)	180.3 (0.0)	185.6 (3.3)
Self-reported race								
Southern European/Mediterranean, %	21.7	18.2	16.5	14.2	29.0	25.5	23.7	21.0
Scandinavian, %	4.4	6.2	7.2	9.0	7.6	9.6	12.6	13.7
Other Caucasian, %	66.2	70.6	71.7	72.6	55.0	61.0	60.5	62.7
Non-white, %	7.7	5.0	4.6	4.2	8.4	3.9	3.1	2.7
Family history of melanoma, %	6.8	7.1	7.4	7.6	4.0	3.8	3.9	3.9
Red/blonde hair, %	13.4	15.1	16.1	17.6	11.1	14.3	14.7	14.9
Presence of arm moles, %	33.7	36.2	37.2	39.2	30.6	32.8	32.7	33.5
Painful burn reaction as a child/adolescent, %	14.0	14.5	14.4	15.7	23.5	24.6	24.6	25.8
Number of blistering sun burn ≥5, %	6.4	7.5	7.4	8.3	31.6	34.8	35.3	37.4
Current smoking, %	32.1	32.8	33.5	35.2	8.7	9.2	9.3	9.4
Alcohol intake (g per day), mean (s.d.)	5.8(10.6)	6.5(11.0)	6.8(11.3)	7.2(11.7)	10.5(15.2)	11.8(16.0)	12.1(16.4)	12.8(17.2)
Body mass index (kg m <sup>-2</sup> ), mean (s.d.)	24.0(4.3)	23.9(4.2)	23.6(4.1)	23.5(4.0)	25.7(3.7)	25.4(3.1)	25.4(3.1)	25.5(3.1)
Physical activity (metabolic-equivalents hours per wk), mean (s.d.)	14.0(20.6)	13.8(20.6)	14.3(21.3)	14.2(21.5)	20.5(29.8)	20.6(28.4)	20.9(30.8)	21.3(30.4)
Menopausal status/PMH status								
Premenopausal, %	80.5	80.5	80.5	80.3				
PMH never use, %	9.2	9.0	9.3	9.0				
PMH current use, %	6.5	6.9	6.8	7.2				
PMH past use, %	3.8	3.7	3.4	3.6				

Abbreviations: HPFS = Health Professionals Follow-up Study; NA = not available; NHS = Nurses' Health Study; PMH = postmenopausal hormone use. Note: Values are means (s.d.) or percentages and are standardised to the age distribution of the study populations. Values of multi-level categorical variables may not sum to 100% due to rounding.  
<sup>a</sup>Value is not age-adjusted.

**Table 2. HRs and 95% CIs for the associations of height (per 10 cm increase) with SCC and BCC risk**

	NHS (Female)		HPFS (Male)		Meta-analysis		
	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value	P for Heterogeneity
<b>SCC</b>							
Model 1	1.18 (1.09, 1.28)	<0.0001	1.18 (1.08, 1.28)	<0.0001	1.18 (1.11, 1.25)	<0.0001	0.99
Model 2	1.14 (1.06, 1.24)	0.0009	1.16 (1.07, 1.26)	0.0006	1.15 (1.09, 1.22)	<0.0001	0.80
Model 3	1.10 (1.02, 1.19)	0.0198	1.10 (1.01, 1.21)	0.0215	1.10 (1.04, 1.17)	0.0011	0.92
Model 4	1.09 (1.00, 1.18)	0.0429	1.09 (1.00, 1.18)	0.0593	1.09 (1.02, 1.15)	0.0057	0.99
<b>BCC</b>							
Model 1	1.18 (1.15, 1.20)	<0.0001	1.14 (1.11, 1.18)	<0.0001	1.16 (1.13, 1.20)	<0.0001	0.10
Model 2	1.16 (1.13, 1.18)	<0.0001	1.13 (1.09, 1.16)	<0.0001	1.14 (1.12, 1.18)	<0.0001	0.16
Model 3	1.12 (1.10, 1.15)	<0.0001	1.10 (1.07, 1.13)	<0.0001	1.11 (1.09, 1.13)	<0.0001	0.28
Model 4	1.11 (1.09, 1.14)	<0.0001	1.08 (1.05, 1.12)	<0.0001	1.10 (1.07, 1.13)	<0.0001	0.20

Abbreviations: BCC = basal cell carcinoma; CI = confidence interval; HPFS = Health Professionals Follow-up Study; HR = hazard ratio; NHS = Nurses' Health Study; PMH = postmenopausal hormone use; SCC = squamous cell carcinoma. (a) Model 1, age-adjusted. (b) Model 2 adjusted for age, smoking status (never, past, current 1–14, 15–24, or 25+ cigarettes/day), alcohol intake (no, <5.0, 5.0–9.9, 10.0–19.9, or 20.0+ g per day), body mass index (<25.0, 25.0–29.9, 30.0–34.9, or 35.0+ kg m<sup>-2</sup>), physical activity (<3.0, 3.0–8.9, 9.0–17.9, 18.0–26.9 or 27.0+ metabolic equivalent hours per week), menopausal status/postmenopausal hormones use (premenopausal, PMH never, PMH past, or PMH current; only in the NHS). (c) Model 3 adjusted for covariates in Model 2, plus natural hair colour (red, blonde, light brown, dark brown, or black), childhood/adolescent sunburn reaction (none or some redness, burn, painful burn or blisters), family history of melanoma (yes or no), number of severe sunburns over life time (never, 1–2 times, 3–5 times, or 6+ times), mole count (none, 1–2, 3–9, 10+), and states lived at birth, age 15, and age 30 (UV index ≤5, =6, or ≥7). (d) Model 4 adjusted for covariates in Model 3 and race (Southern European/Mediterranean, Scandinavian, other Caucasian, or other race group).

risk of SCC and BCC respectively (Supplementary Tables 4 and 5). However, none was significantly associated with risk of skin cancers after Bonferroni correction. Mean values and ranges of the genetic scores combining all 593 independent ( $R^2$  for LD <0.1) height-related SNPs were similar among Illumina, Affy, and Omni data sets (Supplementary Table 3). The genetic scores were significantly associated with height in our genetic data sets ( $P$ -value =  $2.3 \times 10^{-37}$  and  $1.8 \times 10^{-47}$  for simple and weighted genetic scores, respectively). However, we observed no significant association between the scores and risk of NMSC

(Table 4). The results for simple count score and weighted score were similar.

**Sensitivity analysis and validation of self-reported ancestry.** The Pearson correlations between self-reported European ancestry and the first ancestry coordinate were 0.23, 0.28, and 0.31 in Affy, Illumina, and Omni data sets, respectively (all  $P$ -values <0.0001). The Pearson correlations between self-reported European ancestry and the second ancestry coordinate were -0.14, -0.16, and -0.16 in Affy, Illumina, and Omni data sets, respectively (all

P-values <0.0001). Results of the multivariable-adjusted models with self-reported ancestry and the models with genetic ancestry were not materially different (Supplementary Table 6).

**DISCUSSION**

In this analysis of two large and well-characterised cohorts, height was positively associated with risk of both SCC and BCC. To assess confounding due to potential factors, we fitted three multivariable models and gradually added covariates. The magnitude of associations changed the most when skin cancer constitutional factors and sunburns were adjusted for. Self-reported race did not alter the estimates materially when other covariates were already in the models. The multivariable-adjusted HRs for BCC risk among women were greater than the corresponding ones among men, though tests of gender difference did not yield any significant findings. There were much fewer events for SCC than for BCC over the follow-up period; thus the CIs for the former were wider.

Height could represent certain environmental and host factors which may influence the risk of NMSCs. Our results suggest that environmental risk factors such as BMI, smoking, physical activities, pigmentations, sunburn history, and ancestry within the white population are unlikely to explain the observed positive association. Other potential explanations for height-NMSC relation include the link between height and a series of early-life exposures such as nutritional status, living conditions, and serious disease during childhood/adolescence (Silventoinen, 2003). Living conditions and health status in early life could be ruled out, because childhood poverty and illness have been suggested to correlate with an increased risk of subsequent cancers (Smith *et al*, 1998), but are themselves associated with decreased height in adulthood. Thus, stature as a proxy of childhood nutrient status may better explain the positive association between height and cancer. Both animal

studies and epidemiological studies have shown that reduced caloric intake during development reduces future risk of malignancy (Ross and Bras, 1965; Ross and Bras, 1971; Frankel *et al*, 1998). Attention has also been focused on the potential mechanistic relevance of growth factors and hormones. Higher levels of circulating insulin-like growth factor promote linear growth during childhood and have been shown to accelerate cell proliferation (Ish-Shalom *et al*, 1997) and to inhibit apoptosis (Milazzo *et al*, 1992). Another possible explanation is that height may be associated with greater skin surface area, which may put more skin cells at risk of malignant transformation and progression to skin cancer (Albanes and Winick, 1988). However, we were unable to investigate the roles of body surface area or childhood nutrient intake due to unavailability of relevant data.

Genetic factors contribute strongly to adult height. It has been estimated that 80% of the variation in height in Western populations is determined by genetics (McEvoy and Visscher, 2009). Some have proposed that the association between height and cancers may result from shared genetic components. Certain genes linked with height are also related to cancer regulatory pathways such as p53 and HH/PTCH (Tripaldi *et al*, 2013). In addition, height-related SNPs reported by the GIANT consortium have also been associated with risk of testicular cancer and prostate cancer (Wood *et al*, 2014). Yet, it remains unclear whether these height SNPs are tied to skin cancer risk, individually or jointly. In our study, none of the 687 height-related SNPs was significantly associated with SCC or BCC risk after correcting for multiple comparisons. The genetic scores combining all independent SNPs showed no significant association with risk of SCC or BCC. Despite epidemiologic evidence of a link between height and NMSC, we did not detect an association between height GWAS SNPs and NMSC. However, evidence has been insufficient to rule out the roles of genetics. It is possible that the selected SNPs did not cover those genes involved in both height and skin cancers, as we only included height SNPs reported by GWAS studies. Furthermore, regulations beyond the gene level and gene-environment interactions may also explain the observed association.

The strengths of the current study include prospective design with long-term follow-up and high follow-up rate, availability of detailed information on a wide variety of covariates, involvement of both women and men, and targeting on SCC and BCC separately. A major advantage is that we examined the associations between height and skin cancers more thoroughly than what has previously been reported. Potential confounding factors, such as pigmentation and sunburn history, which are critical for skin cancers and have not been considered before, were included in our Cox models. We also considered interactions between height and other covariates. In sensitivity analysis, ancestry within the white population was assessed directly using genetic data, though adjustment for genetic ancestry did not change the results materially. This may be due to lack of power in the genetic subsets and/or the control of skin cancer constitutional factors

**Table 3. Number of NMSC cases and controls<sup>a</sup> in each of the combined data sets**

Data set	BCC		SCC	
	cases	controls	cases	controls
Affymetrix	1781	4304	247	4500
Illumina	1055	1929	134	2029
OmniExpress	1062	2297	146	2433
Total	3898	8530	527	8962

Abbreviations: BCC = basal cell carcinoma; HPFS = Health Professionals Follow-up Study; NHS = Nurses' Health Study; NMSC = non-melanoma skin cancer; SCC = squamous cell carcinoma.

<sup>a</sup>1976 and 1986 were considered baseline years for the NHS and the HPFS respectively. Baseline cancer cases were excluded. Skin cancer cases who had diagnosis of other common cancers before diagnosis of skin cancers were excluded; controls who had other cancers were excluded.

**Table 4. Associations between genetic scores of height-related SNPs and risk of NMSC<sup>a</sup>**

		Illumina			Affymetrix			OmniExpress			Meta-analysis <sup>b</sup>		
		Beta	s.e.	P-value	Beta	s.e.	P-value	Beta	s.e.	P-value	Beta	P-value	P for heterogeneity
Simple count genetic score	SCC	0.001	0.003	0.784	0.002	0.002	0.473	0.003	0.003	0.231	0.002	0.169	0.895
	BCC	0.002	0.001	0.206	-0.0001	0.001	0.887	0.001	0.001	0.499	0.001	0.094	0.332
Weighted genetic score	SCC	0.001	0.003	0.815	0.002	0.002	0.454	0.004	0.003	0.120	0.002	0.125	0.767
	BCC	0.001	0.001	0.295	2.008E-05	0.001	0.9834	0.001	0.001	0.400	0.001	0.244	0.726

Abbreviations: BCC = basal cell carcinoma; NMSC = non-melanoma skin cancer; SCC = squamous cell carcinoma; SNP = single-nucleotide polymorphism.

<sup>a</sup>Logistic regression models were used to assess the relationship between genetic score of height SNPs and risk of skin cancers, adjusting for age, gender, and the top three principal components.

<sup>b</sup>Analyses were first conducted within each of the platform-specific genetic data sets. We used fixed-effect meta-analysis to obtain a combined estimate.

which have already partly explained variation of ancestry in the models. Moreover, our novel analysis of the associations between height-related genetic variants and risk of skin cancers may eventually yield a better understanding of the underlying mechanisms. To our best knowledge, no such analysis has been conducted for skin cancers.

We also acknowledge several potential limitations of the present study. First, height was self-reported rather than measured in our cohorts, which could result in misclassification. However, any misclassification would be non-differential with respect to disease occurrence, because information on height was collected prior to the development of skin cancers. Because non-differential misclassification would bias the estimation downwards, that could not account for the observed positive association. Second, BCC cases were self-reported without further pathological confirmation. However, the high validity of self-reported BCC in these medically sophisticated populations has been confirmed in previous studies (Colditz *et al*, 1986). In addition, using the self-reported BCC cases, our group identified the previously well-documented genetic variant in the MC1R gene as the top risk locus in our GWAS for BCC (Nan *et al*, 2011). These data support the validity of self-report of BCC in our study. Third, we did not have information on all relevant confounding variables. For example, data on socioeconomic status, which might affect both height and cancer incidence, were not available. However, our study used cohorts of health-care providers, which has the advantage of minimising confounding by educational attainment and adult socioeconomic status. In addition, adjustment for socioeconomic factors did not affect risk estimates for association between height and cancer in previous large studies (Sung *et al*, 2009; Green *et al*, 2011; Kabat *et al*, 2013b). We also lacked information on childhood nutritional status, for which height may be a marker. Finally, our cohorts consist primarily of white health professionals and thus results may not be generalisable. However, such homogeneity in a study population would minimise confounding by socioeconomic status and differential access to health care and assure a high quality of returned data.

In conclusion, our data from two large cohorts provide further evidence that greater height is associated with increased risk of SCC and BCC. These associations were not explained by confounding by known risk factors, nor were they modified by those risk factors. No significant association was observed between height-related genetic variants and risk of NMSC, whether individually or jointly. More functional and epidemiological studies on height-related SNPs are needed to confirm our findings. Additional research involving a range of pre-adult exposures, such as diet, psychosocial stress, chronic illness, and social circumstances, which are rarely directly measured in existing data sets, may help clarify possible mechanisms underlying the positive associations.

## ACKNOWLEDGEMENTS

We are indebted to the participants in the NHS and HPFS for their dedication to this research. We thank the following state cancer registries for their help: Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Nebraska, New Hampshire, New Jersey, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Virginia, Washington, and Wyoming. The authors assume full responsibility for analyses and interpretation of these data. This work is supported by NIH R01 CA49449, P01 CA87969, UMI CA186107, and UMI CA167552.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

- Ainsworth BE, Haskell WL, Leon AS, Jacobs Jr DR, Montoye HJ, Sallis JF, Paffenbarger Jr RS (1993) Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc* **25**(1): 71–80.
- Albanes D, Winick M (1988) Are cell number and cell proliferation risk factors for cancer? 1. *J Natl Cancer Inst* **80**(10): 772–775.
- Armstrong BK, Krickler A, English DR (1997) Sun exposure and skin cancer. *Australas J Dermatol* **38**(Suppl 1): S1–S6.
- Balding DJ (2006) A tutorial on statistical methods for population association studies. *Nat Rev Genet* **7**(10): 781–791.
- Cavelaars A, Kunst A, Geurts J, Crialesi R, Grötvedt L, Helmert U, Lahelma E, Lundberg O, Mielck A, Rasmussen NK (2000) Persistent variations in average height between countries and between socio-economic groups: an overview of 10 European countries. *Ann Hum Biol* **27**(4): 407–421.
- Chasan-Taber S, Rimm EB, Stampfer MJ, Spiegelman D, Colditz GA, Giovannucci E, Ascherio A, Willett WC (1996) Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology* **7**(1): 81–86.
- Colditz GA, Martin P, Stampfer MJ, Willett WC, Sampson L, Rosner B, Hennekens CH, Speizer FE (1986) Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am J Epidemiol* **123**(5): 894–900.
- Cutler C, Foulkes W, Brunet J, Flanders T, Shibata H, Narod S (1996) Cutaneous malignant melanoma in women is uncommonly associated with a family history of melanoma in first-degree relatives: a case-control study. *Melanoma Res* **6**(6): 435–440.
- D'Arcy C, Holman J, Armstrong BK (1984) Pigmentary traits, ethnic origin, benign nevi, and family history as risk factors for cutaneous malignant melanoma. *J Natl Cancer Inst* **72**(2): 257–266.
- English DR, Armstrong BK, Krickler A, Winter MG, Heenan PJ, Randell PL (1998) Demographic characteristics, pigmentary and cutaneous risk factors for squamous cell carcinoma of the skin: a case-control study. *Int J Cancer* **76**(5): 628–634.
- Frankel S, Gunnell DJ, Peters TJ, Maynard M, Smith GD (1998) Childhood energy intake and adult mortality from cancer: the Boyd Orr Cohort Study. *Bmj* **316**(7130): 499–504.
- Gallagher RP, Hill GB, Bajdik CD, Coldman AJ, Fincham S, McLean DI, Threlfall WJ (1995a) Sunlight exposure, pigmentation factors, and risk of nonmelanocytic skin cancer: II. Squamous cell carcinoma. *Arch Dermatol* **131**(2): 164–169.
- Gallagher RP, Hill GB, Bajdik CD, Fincham S, Coldman AJ, McLean DI, Threlfall WJ (1995b) Sunlight exposure, pigmentary factors, and risk of nonmelanocytic skin cancer: I. Basal cell carcinoma. *Arch Dermatol* **131**(2): 157–163.
- Gallus S, Naldi L, Martin L, Martinelli M (2006) Anthropometric measures and risk of cutaneous malignant melanoma: a case-control study from Italy. *Melanoma Res* **16**(1): 83–87.
- Grasgruber P, Cacek J, Kalina T, Sebera M (2014) The role of nutrition and genetics as key determinants of the positive height trend. *Econ Hum Biol* **15**: 81–100.
- Green J, Cairns BJ, Casabonne D, Wright FL, Reeves G, Beral V. collaborators MWS (2011) Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. *Lancet Oncol* **12**(8): 785–794.
- Han J, Colditz GA, Hunter DJ (2006) Risk factors for skin cancers: a nested case-control study within the Nurses' Health Study. *Int J Epidemiol* **35**(6): 1514–1521.
- Hunter DJ, Colditz GA, Stampfer MJ, Rosner B, Willett WC, Speizer FE (1990) Risk factors for basal cell carcinoma in a prospective cohort of women. *Ann Epidemiol* **1**(1): 13–23.
- Ish-Shalom D, Christoffersen C, Vorwerk P, Sacerdoti-Sierra N, Shymko R, Naor D, De Meyts P (1997) Mitogenic properties of insulin and insulin analogues mediated by the insulin receptor. *Diabetologia* **40**(2): S25–S31.

- Kabat GC, Anderson ML, Heo M, Hosgood HD, Kamensky V, Bea JW, Hou L, Lane DS, Wactawski-Wende J, Manson JE (2013a) Adult stature and risk of cancer at different anatomic sites in a cohort of postmenopausal women. *Cancer Epidemiol Biomarkers Prevent* **22**(8): 1353–1363.
- Kabat GC, Heo M, Kamensky V, Miller AB, Rohan TE (2013b) Adult height in relation to risk of cancer in a cohort of Canadian women. *Int J Cancer* **132**(5): 1125–1132.
- McEvoy BP, Visscher PM (2009) Genetics of human height. *Econ Hum Biol* **7**(3): 294–306.
- Milazzo G, Giorgino F, Damante G, Sung C, Stampfer MR, Vigneri R, Goldfine ID, Belfiore A (1992) Insulin receptor expression and function in human breast cancer cell lines. *Cancer Res* **52**(14): 3924–3930.
- Nan H, Xu M, Kraft P, Qureshi AA, Chen C, Guo Q, Hu FB, Curhan G, Amos CI, Wang L-E (2011) Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* **20**(18): 3718–3724.
- Narayanan DL, Saladi RN, Fox JL (2010) Review: ultraviolet radiation and skin cancer. *Int J Dermatol* **49**(9): 978–986.
- Olsen CM, Green AC, Zens MS, Stukel TA, Bataille V, Berwick M, Elwood JM, Gallagher R, Holly EA, Kirkpatrick C (2008) Anthropometric factors and risk of melanoma in women: a pooled analysis. *Int J Cancer* **122**(5): 1100–1108.
- Rehnan AG (2011) Height and cancer: consistent links, but mechanisms unclear. *Lancet Oncol* **12**(8): 716–717.
- Rogers HW, Weinstock MA, Harris AR, Hinckley MR, Feldman SR, Fleischer AB, Coldiron BM (2010) Incidence estimate of nonmelanoma skin cancer in the United States, 2006. *Arch Dermatol* **146**(3): 283–287.
- Ross MH, Bras G (1965) Tumor incidence patterns and nutrition in the rat. *J Nutr* **87**(3): 245–260.
- Ross MH, Bras G (1971) Lasting influence of early caloric restriction on prevalence of neoplasms in the rat. *J Natl Cancer Inst* **47**(5): 1095–1114.
- Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, Willett WC (1989) Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int J Epidemiol* **18**(4): 858–867.
- Shors AR, Solomon C, McTiernan A, White E (2001) Melanoma risk in relation to height, weight, and exercise (United States). *Cancer Causes Control* **12**(7): 599–606.
- Silventoinen K (2003) Determinants of variation in adult body height. *J Biosoc Sci* **35**(02): 263–285.
- Smith GD, Hart C, Blane D, Hole D (1998) Adverse socioeconomic conditions in childhood and cause specific adult mortality: prospective observational study. *Bmj* **316**(7145): 1631–1635.
- Sung J, Song Y-M, Lawlor DA, Smith GD, Ebrahim S (2009) Height and site-specific cancer risk: a cohort study of a Korean adult population. *Am J Epidemiol* **170**(1): 53–64.
- Thune I, Olsen A, Albrektsen G, Tretli S (1993) Cutaneous malignant melanoma: association with height, weight and body-surface area. A prospective study in Norway. *Int J Cancer* **55**(4): 555–561.
- Tripaldi R, Stuppia L, Alberti S (2013) Human height genes and cancer. *Biochim Biophys Acta* **1836**(1): 27–41.
- Tung RC, Vidimos AT (2002) Non-melanoma skin cancer. Retrieved March 29: 2007.
- Turchin MC, Chiang CW, Palmer CD, Sankararaman S, Reich D, Hirschhorn JN, Consortium GloAT (2012) Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat Genet* **44**(9): 1015–1019.
- Wang C, Zhan X, Bragg-Gresham J, Kang HM, Stambolian D, Chew EY, Branham KE, Heckenlively J, Study TF, Fulton R (2014) Ancestry estimation and control of population stratification for sequence-based association studies. *Nat Genet* **46**(4): 409–415.
- Wang C, Zhan X, Liang L, Abecasis GR, Lin X (2015) Improved ancestry estimation for both genotyping and sequencing data using projection procrustes analysis and genotype imputation. *Am J Hum Genet* **96**(6): 926–937.
- Willett WC, Sampson L, Browne ML, Stampfer MJ, Rosner B, Hennekens CH, Speizer FE (1988) The use of a self-administered questionnaire to assess diet four years in the past. *Am J Epidemiol* **127**(1): 188–199.
- Wirén S, Häggström C, Ulmer H, Manjer J, Bjørge T, Nagel G, Johansen D, Hallmans G, Engeland A, Concin H (2014) Pooled cohort study on height and risk of cancer and cancer death. *Cancer Causes Control* **25**(2): 151–159.
- Wolf AM, Hunter DJ, Colditz GA, Manson JE, Stampfer MJ, Corsano KA, Rosner B, Kriska A, Willett WC (1994) Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol* **23**(5): 991–999.
- Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan JA, Kutalik Z (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**(11): 1173–1186.
- Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**(7): 565–569.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)