

# Accuracy of Self-Reported Nevus and Pigmentation Phenotype Compared with Clinical Assessment in a Population-Based Study of Young Australian Adults

Anne E. Cust<sup>1</sup>, Kristen M. Pickles<sup>2</sup>, Chris Goumas<sup>1</sup>, Thao Vu<sup>1</sup>, Helen Schmid<sup>3</sup>, Eduardo Nagore<sup>4</sup>, John Kelly<sup>5</sup>, Joanne F. Aitken<sup>6</sup>, Graham G. Giles<sup>7,8</sup>, John L. Hopper<sup>7</sup>, Mark A. Jenkins<sup>7</sup>, and Graham J. Mann<sup>3</sup>

## Abstract

**Background:** Awareness of individual risk may encourage improved prevention and early detection of melanoma.

**Methods:** We evaluated the accuracy of self-reported pigmentation and nevus phenotype compared with clinical assessment, and examined agreement between nevus counts from selected anatomical regions. The sample included 456 cases with invasive cutaneous melanoma diagnosed between ages 18 to 39 years and 538 controls from the population-based Australian Melanoma Family Study. Participants completed a questionnaire about their pigmentation and nevus phenotype, and attended a dermatologic skin examination.

**Results:** There was strong agreement between self-reported and clinical assessment of eye color [ $\kappa$ , = 0.78; 95% confidence interval (CI), 0.74–0.81]; and moderate agreement for hair color ( $\kappa$  = 0.46; 95% CI, 0.42–0.50). Agreement between self-reported skin color and spectrophotometer-derived measurements was

poor ( $\kappa$  = 0.12; 95% CI, 0.08–0.16) to moderate (Spearman correlation  $r_s$  = –0.37; 95% CI, –0.32 to –0.42). Participants tended to underestimate their nevus counts and pigmentation; men were more likely to underreport their skin color. The  $r_s$  was 0.43 (95% CI, 0.38–0.49) comparing clinical total body nevus counts with self-reported nevus categories. There was good agreement between total body nevus counts and site-specific nevus counts, particularly on both arms.

**Conclusions:** Young adults have suboptimal accuracy when assessing important risk characteristics including nevus numbers and pigmentation. Measuring nevus count on the arms is a good predictor of full body nevus count.

**Impact:** These results have implications for the likely success of targeted public health programs that rely on self-assessment of these factors. *Cancer Epidemiol Biomarkers Prev*; 24(4); 736–43. ©2015 AACR.

## Introduction

Melanoma is a common malignancy in fair-skinned populations (1–3). In Australia, melanoma is the most frequently diagnosed cancer among those ages 15 to 44 years, accounting for 23%

of incident cancers and 8% of cancer deaths in this age group (4). Melanoma is a highly preventable cancer (5, 6) and has an excellent prognosis when detected early (7). There is no formal melanoma screening program in the general population, so it is particularly important that individuals are able to accurately assess key melanoma risk factors so that those at high risk can identify themselves as such and take appropriate preventative actions, such as increasing sun protection and having regular skin checks, from a young age.

Visibly identifiable risk factors for melanoma include nevi (8) and pigmentation indicative of a sun-sensitive phenotype such as fair skin, red or blonde hair, and blue or green eye color (9). For assessing the validity of self-reported phenotypic characteristics, the "gold standard" measure is clinical assessment by dermatologists, who have expertise in classifying level of pigmentation and identifying nevi.

There are several limitations to previous validation studies of self-reported phenotype characteristics. Although young adults are considered a key target group for melanoma prevention campaigns, most previous validation studies have been conducted among older adults. Many previous studies have also used convenience or clinic-based sampling, and had small sample sizes. It is also uncertain to what degree people with and without melanoma differ in the accuracy of their self-reported phenotype, and whether this has the potential to bias risk estimates derived from

<sup>1</sup>Cancer Epidemiology and Services Research (CESR), Sydney School of Public Health, The University of Sydney, Sydney, Australia. <sup>2</sup>Sydney School of Public Health, The University of Sydney, Sydney, Australia. <sup>3</sup>Centre for Cancer Research, University of Sydney at Westmead Millennium Institute for Medical Research and Melanoma Institute Australia, Sydney, Australia. <sup>4</sup>Department of Dermatology, Instituto Valenciano de Oncología, Valencia, Spain. <sup>5</sup>Victorian Melanoma Service, Alfred Hospital, Melbourne, Australia. <sup>6</sup>Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Spring Hill, Brisbane, Australia. <sup>7</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia. <sup>8</sup>Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia.

**Note:** Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

**Corresponding Author:** Anne E. Cust, Cancer Epidemiology and Services Research (CESR), Level 6, The Lifehouse, 119-143 Missenden Road, Camperdown NSW 2050, Australia. Phone: 61-2-8627-1565; Fax: 61-2-9036-5292; E-mail: [anne.cust@sydney.edu.au](mailto:anne.cust@sydney.edu.au)

doi: 10.1158/1055-9965.EPI-14-1203

©2015 American Association for Cancer Research.

case-control studies of melanoma. Several nevus validation studies have focused on assessment of larger and atypical nevi (10, 11); however, the total number of common melanocytic nevi on the body is a stronger risk factor for melanoma (8).

The aim of this study was to evaluate the accuracy of self-reported pigmentation and nevus phenotype compared with more objective measures recorded by dermatology trainees during a clinical skin examination, for 994 young adults ages 18 to 44 years including 456 people with melanoma (cases) and 538 people without melanoma (controls) who participated in the Australian Melanoma Family Study. We also examined predictors of disagreement. A secondary aim was to examine agreement between clinically measured nevus counts on the whole body to that from selected anatomical regions, to strengthen the evidence for using localized nevus count as a proxy measure of whole body nevus counts.

## Materials and Methods

### Study sample

The Australian Melanoma Family Study is a multicentre, population-based, case-control-family study of invasive cutaneous melanoma diagnosed between ages 18 and 39 years. The study design, recruitment, data collection, and participant characteristics have been previously described in detail (12). Recruitment of case and control participants was locally conducted in Sydney, Melbourne, and Brisbane, Australia. Approval for the study was obtained from the ethics committees of the three coordinating centers and the cancer registries. All participants provided written, informed consent.

**Case participants.** Cases were identified from population-based state cancer registries, diagnosed between July 1, 2000 and December 31, 2002 at ages 18 to 39 years with incident, histopathologically confirmed, first-primary invasive cutaneous melanoma. Participation was 54% of those eligible and 76% of those contactable.

**Control participants: population, spouse/friend, and sibling.** Population controls were ages between 18 years and 39 years at the time of approach and had no history of invasive or *in situ* melanoma. They were selected from the electoral roll (registration to vote is compulsory for Australian citizens ages 18 years and over) and were frequency matched to cases by city, age (within 5 years), and sex. Participation was 23% of those eligible and 42% of those contactable. Eligible spouse or friend controls were a spouse, partner, or friend nominated by a case as a potential control subject. They were eligible if they were at least 18 years of age and had no history of invasive or *in situ* melanoma; there were no other age, sex, or residency restrictions. A spouse or friend was nominated as a potential control subject by 59% of cases, and participation was 80% of those nominated. We sought to recruit all siblings of cases for whom the case permitted contact. Unaffected siblings of cases were considered as sibling controls, and at least one sibling control was identified for 91% of cases.

### Self-reported pigmentation phenotype and nevus counts

Participants completed a questionnaire in which they reported their skin color (very fair, fair, olive or brown, Asian, black, other), eye color (blue or gray, green or hazel, brown or black), natural hair color at age 18 years (red or auburn, fair or blonde, light or mouse brown, gray, dark brown, black), ability to tan (tanning

response to repeated sun exposure in summer: deep tan, moderately tan, mild tan, no tan), propensity to burn (response to prolonged sun exposure in summer: severe sunburn with blistering, painful sunburn with peeling, mild sunburn, no sunburn), usual tanning and sunburn response to prolonged or repeated exposure of skin to sunlight in summer, the number of nevi covering their body (described pictorially as none, few, some, many), and were asked to have someone count the number of all moles on their back (using picture guides to define the area and describe moles; ref. 12).

### Clinical assessment of pigmentation and nevus phenotype

A clinical skin examination was completed by 73% of cases, 55% of population controls, 67% of spouse or friend controls, and 43% of sibling controls. Skin examinations were conducted at dermatology clinics in Brisbane, Sydney, and Melbourne by dermatology trainees trained on the study protocol. A dermatologist with experience in using this protocol trained examiners in classification of lesions to minimize undercounting. Participants removed their clothing except for underpants and bra. Measurement of nevi was based on international guidelines (13). Separate counts were made for melanocytic nevi of 2 to 5 mm and >5 mm, raised nevi of >2 mm, and clinically atypical nevi of >2 mm, on 30 body sites (Supplementary Fig. S1). Only nevi over 2 mm were considered, to minimize confusion with freckles and lentigines. Natural hair color at age 18 years (red, fair or blonde, brown, black) and eye color (light blue or blue or blue/gray, green or hazel, light brown or brown or dark brown) were recorded by using wig hair swatches and eye photographs.

**Skin reflectance.** Reflected skin color, a correlate of melanin content (14, 15), was objectively recorded (mean of 6 readings) from both the outer and inner part of the subject's left upper arm using a BYK-Gardner CGSS hand-held reflectance spectrophotometer with standard reflectance at 685 nm, calibrated before each session. The multi-wavelength data quantify color using the Commission Internationale de l'Éclairage  $L^* a^* b^*$  color space parameters: with dimension  $L^*$  indicating lightness of color (smaller values indicate darker color),  $a^*$  the position between red and green, and  $b^*$  the position between yellow and blue (smaller values indicate lighter color; ref. 16). Inner arm  $L^*$  values have been used to describe natural skin color,  $b^*$  values to describe tanning, and  $a^*$  values to describe erythema (14, 16–18). Spectrophotometer readings from the inner arm were also converted to individual typography angle (ITA) scores, calculated using the values of  $L^*$  and  $b^*$  according to the formula:  $ITA^\circ = (\text{ArcTAN}((L^* - 50)/b^*)) * 180/3.14159$  (19), where Arc Tangent is an inverse trigonometric function expressed in radians. ITA scores have been shown to correlate with melanin content (19) and have been used to classify skin color into six categories (19, 20); we collapsed to three categories using cutoff points: very fair  $\geq 55^\circ$  > fair  $\geq 28^\circ$  > olive/brown/dark (19).

### Statistical analysis

Data were available for 994 participants ages 18 to 44 years (456 cases and 538 controls, of which 131 population-based controls, 168 spouse, partner or friend controls and 239 sibling controls), who had completed a self-report questionnaire and attended a clinical skin examination.

We compared self-reported and clinical measures using Cohen's kappa ( $\kappa$ ) and weighted kappa ( $\kappa_w$ ) statistics for categorical and ordinal variables (21) and intraclass correlation

coefficients (ICC) for continuous variables. Kappa statistics are interpreted as: <0.20 poor agreement beyond chance, 0.21 to 0.40 fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 good agreement, and 0.81 to 1.00 excellent agreement (22). Box-and-whisker plots and Spearman correlation coefficients ( $r_s$ ) were used to compare ordinal variables with continuous variables, for example, when comparing self-reported nevus categories (none, few, some, many) with clinically measured nevus counts. Spectrophotometer ( $L^* a^* b^*$ ) values and nevus counts were analyzed as continuous variables, and as quartiles based on the distribution in controls.

As there were slight differences in the categories for eye and hair color measures between self-report and clinical assessment, we harmonized the categories for analysis. We used four categories for natural hair color at the age of 18 years: red (self-report: red or auburn; clinical: red), blonde/fair (self-report: fair or blonde, gray; clinical: fair or blonde), brown (self-report: light or mouse brown, dark brown; clinical: brown), and black (self-report: black; clinical: black). We used three eye color categories: blue/gray (self-report: blue or gray; clinical: light blue or blue or blue/gray), Green/hazel (self-report: green or hazel; clinical: green or hazel), and brown/black (self-report: brown or black; clinical: light brown or brown or dark brown). Self-reported skin color was compared with spectrophotometer values as specified above; due to small numbers, we grouped participants with olive or brown, Asian, black, or other skin color as one category "olive/dark."

We examined agreement between clinically measured nevus counts on the whole body to those from selected anatomical regions, including the posterior body (from head to toe on the back side of the body), the back, both arms, the right arm, and both thighs (see Supplementary Fig. S1 for classification of body sites). We examined whether any personal characteristics or other factors were associated with disagreement between self-reported and clinically measured pigmentation phenotype variables (hair, skin, and eye color), using log-binomial regression models to estimate probability ratios and 95% confidence intervals (CI) after adjusting for age, sex, and city of recruitment.

All participants (cases and different control groups) were combined for the main analysis, and we conducted subgroup analyses to examine whether the accuracy of self-reported measures differed by case-control status and sex. All analyses were performed using SAS Statistical Software (version 9.3, SAS Institute), and statistical significance was inferred at two-sided  $P < 0.05$ .

## Results

Demographic and phenotype characteristics of cases and controls are shown in Table 1. Nearly all participants were of white European descent and 42% of both cases and controls had a University degree. Cases were more likely to have a sun-sensitive phenotype and more nevi than controls.

### Hair, eye, and skin color categories

There was excellent agreement between self-reported and clinical assessment of eye color ( $\kappa = 0.78$ ); and moderate agreement for hair color ( $\kappa = 0.46$ ; Table 2). Agreement between self-reported skin color and spectrophotometer-derived measurements was weak ( $\kappa = 0.12$ ) when using the prespecified ITA category cutoff points. Participants tended to underreport

**Table 1.** Demographic and clinically measured phenotype characteristics of cases and controls

Characteristic	Cases (n = 456) n (%)	Controls (n = 538) n (%)
Sex		
Male	170 (37)	232 (43)
Female	286 (63)	306 (57)
Age at diagnosis/interview (years) <sup>a</sup>		
18–29	139 (30)	120 (22)
30–44	317 (70)	418 (78)
Ethnicity		
Caucasian	443 (97)	516 (96)
Other	13 (3)	22 (4)
Highest level of education		
Junior secondary school	61 (13)	54 (10)
Senior secondary school	116 (26)	124 (23)
Vocational/certificate	86 (19)	134 (25)
University	189 (42)	223 (42)
Hair color <sup>b</sup>		
Red	82 (18)	49 (9)
Blonde/fair	197 (43)	213 (40)
Black/brown	177 (39)	274 (51)
Eye color <sup>b</sup>		
Blue/gray	233 (51)	262 (49)
Green/hazel	133 (29)	155 (29)
Brown/black	87 (19)	120 (22)
Skin color <sup>b,c</sup>		
Very fair	288 (63)	289 (54)
Fair	149 (33)	202 (38)
Olive/brown/dark	18 (4)	43 (8)
Total body nevi $\geq 2 \text{ mm}^2$ (median, 25th–75th centile)	201 (107–307)	91 (37–191)

<sup>a</sup>All cases were <40 years at diagnosis and all population controls were <40 years when ascertained. Cases and controls could be up to age 44 years at interview for this analysis.

<sup>b</sup>Assessed from the clinical skin examinations.

<sup>c</sup>Skin color based on ITA scores, calculated using the values of  $L^*$  and  $b^*$  measured by skin reflectance and based on cutoff points: very fair  $\geq 55^{\circ}$  > fair  $\geq 28^{\circ}$  > olive/brown/dark (19).

blonde/fair and red hair, very fair skin, and blue/gray eyes, compared with clinical assessment.

### Predictors of disagreement for hair, eye, and skin color

There were no significant differences in  $\kappa$  values for hair, eye, or skin color when stratified by case-control status. Men and women had similar agreement for eye and hair color, but skin color agreement appeared slightly higher for women ( $\kappa = 0.17$ ) than for men ( $\kappa = 0.07$ ). We further examined predictors of disagreement in a log-binomial regression model adjusted for age, sex, and city of recruitment (Supplementary Table S1). Again, we observed less disagreement between self-reported and spectrophotometer-measured skin color for women compared with men (probability ratio 0.81; 95% CI, 0.72–0.90). Agreement for eye color ( $P_{\text{difference}} = 0.01$ ) and, to a lesser degree skin color ( $P_{\text{difference}} = 0.05$ ), was worse for participants from Brisbane (lower latitude) than for those living in Sydney and Melbourne (higher latitudes). Age, ethnicity, country of birth, residence-based socioeconomic disadvantage level, education, case-control status, family history of melanoma in a first-degree relative, and nevus counts were not significantly associated with disagreement (Supplementary Table S1). Disagreement for skin color was not a predictor of disagreement for eye or hair color, and vice versa.

**Table 2.** Self-report versus clinical assessment of hair, eye, and skin color

Self-reported	Clinically measured				Measure of agreement	Statistic
Hair color	Red	Blonde/fair	Brown	Black		
Red	73	7	3	0	Percent agreement	66%
Blonde/fair	16	158	3	0	$\kappa$ (95% CI)	0.46 (0.42–0.50)
Brown	41	244	399	8	$\kappa_w$ (95% CI)	0.52 (0.47–0.56)
Black	0	0	10	28		
Eye color	Blue/gray	Green/hazel	Brown/black			
Blue/gray	423	14	0	Percent agreement	86%	
Green/hazel	67	246	33	$\kappa$ (95% CI)	0.78 (0.74–0.81)	
Brown/black	1	23	174	$\kappa_w$ (95% CI)	0.83 (0.80–0.86)	
Skin color <sup>a</sup>	Very fair	Fair	Olive/dark			
Very fair	148	24	4	Percent agreement	43%	
Fair	390	250	29	$\kappa$ (95% CI)	0.12 (0.08–0.16)	
Olive/dark	34	73	28	$\kappa_w$ (95% CI)	0.18 (0.14–0.22)	

NOTE: The gray-shaded boxes indicate perfect agreement between the self-reported and clinical assessment of phenotype; the boxes surrounded by dashed lines indicate that the participant overestimated their pigmentation phenotype; and the boxes surrounded by straight lines indicate the participant underestimated their pigmentation phenotype.

<sup>a</sup>Clinically measured skin color was based on reflectance spectrophotometry L\* and b\* values converted to ITA scores, and categorized according to published ITA cutoff points: very fair  $\geq 55$ ; Fair  $\geq 28$ ; olive/brown/dark (19).

**Skin reflectance correlations with self-reported skin color**

Table 3 shows descriptive statistics and correlations for skin reflectance values (ITA scores, L\*, b\* and a\* scale values) on the upper inner arm compared with self-reported skin color categories. Self-reported skin color was negatively correlated with ITA scores and L\* scale values, and positively correlated with b\* and a\* scale values; all consistent with increasing pigmentation being objectively detected in the skin for people self-reporting darker-colored skin. The strength of the correlation was moderate for ITA scores ( $r_s -0.37$ ), b\* scale values ( $r_s 0.34$ ), and L\* scale values ( $r_s -0.31$ ) and weak for a\* scale values ( $r_s 0.17$ ). Correlations were similar or slightly lower when the outer arm values were used instead of the inner arm values (data not shown). The difference in b\* values between the inner and outer arm measurements, which has been used as an indicator of the degree of tanning (18), was not correlated with skin color ( $r_s 0.00$ ) in our sample.

As b\* values have sometimes been used as an indicator of tanning, and a\* values for erythema, we also compared these skin reflectance values with self-reported ability to tan and propensity to burn, respectively. The  $r_s$  for self-reported ability to tan was 0.41 (95% CI, 0.36–0.46)  $P < 0.001$  for inner arm b\* values and 0.33 (95% CI, 0.27–0.38;  $P < 0.001$ ) for outer arm b\* values. There was poor correlation between a\* values and self-reported propensity to burn:  $r_s = 0.13$  (95% CI, 0.07–0.19;  $P < 0.001$ ) for inner arm values and 0.04 (95% CI, -0.02 to 0.11;  $P = 0.17$ ) for outer arm values.

**Nevus counts**

Figure 1 box-and-whisker plot shows moderate agreement between clinical total body nevus counts and self-reported nevus

density (none, few, some, many) categories ( $r_s = 0.43$ ; 95% CI, 0.38–0.49). There was better distinction between the "some" and "many" categories than for the "none" and "few" categories: the median clinically measured nevus counts for each self-reported category were 50 (none), 86 (few), 170 (some), and 254 (many). There were no significant differences by case–control status or sex.

When self-reported counts of moles on the back were compared with clinically measured nevus counts on the back, the ICC was 0.36 (95% CI, 0.30–0.41; Fig. 2). When nevus counts on the back were categorized into quartiles and compared, the  $\kappa_w$  was 0.28 (95% CI, 0.23–0.32). Agreement for nevus counts on the back did not significantly differ by case–control status or sex. On average, participants underreported their nevus counts by more than half: the median number of nevi recorded on the back by self-assessment was 13 for cases and eight for controls, whereas clinicians recorded a median 38 nevi for cases and 19 for controls.

**Agreement between clinically measured nevus counts on selected anatomical sites**

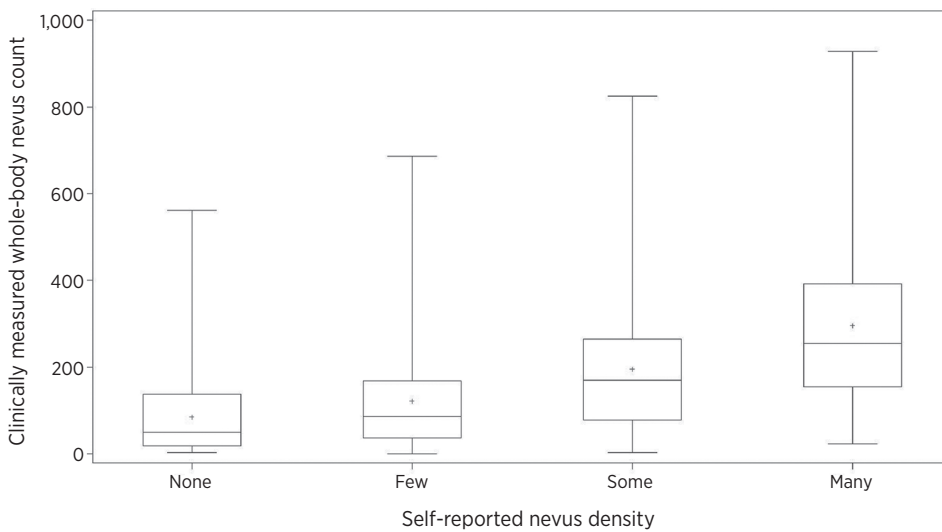
There was generally good agreement when clinically measured total body nevus counts were compared with site-specific nevus counts, particularly for the whole back of the body and both arms (Table 4).

**Discussion**

An important finding of this study is that young adults living in a country with relatively high levels of UV radiation and

**Table 3.** Self-reported skin color compared with objectively measured skin reflectance by spectrophotometry

Self-reported skin color	Objectively measured skin reflectance on the upper inner arm							
	ITA score		L* scale		b* scale		a* scale	
	Median	25th–75th centile	Median	25th–75th centile	Median	25th–75th centile	Median	25th–75th centile
Very fair	60.1	56.5–63.4	69.2	68.1–70.4	11.0	10.0–12.2	6.0	5.3–7.0
Fair	56.1	51.4–60.6	68.0	66.3–69.4	11.9	10.7–13.4	6.6	5.6–7.5
Olive/dark	49.9	42.2–55.2	66.4	63.9–68.2	14.0	11.9–15.8	7.1	5.9–8.3
Spearman $r_s$	-0.37 (-0.32 to -0.42), $P < 0.001$		-0.31 (-0.26 to -0.37), $P < 0.001$		0.34 (0.28–0.39), $P < 0.001$		0.17 (0.11–0.23), $P < 0.001$	



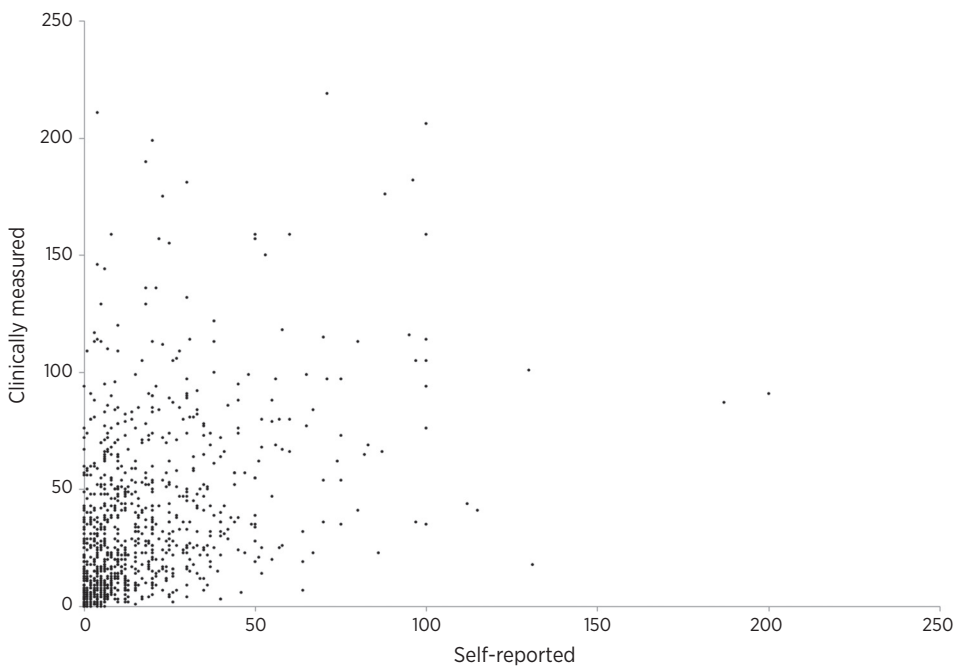
**Figure 1.** Box-and-whisker plot showing agreement between clinically measured whole-body nevus counts (2+ mm) and self-reported nevus density descriptive (none, few, some, many) categories (Spearman  $r_s = 0.43$ ; 95% CI, 0.38–0.49;  $P < 0.001$ ).

elevated melanoma incidence rates tend to underreport their nevus counts and skin, hair, and eye pigmentation phenotype. Young people are a key target group for preventive efforts as they tend to have excessive UV exposure and slow adoption of sun-protective behaviors (20, 23, 24). Awareness of individual risk may encourage improved prevention and early detection of melanoma; however, these results show that young people have suboptimal accuracy when assessing important risk characteristics, including nevus numbers and pigmentation. This has implications for the likely success of targeted public health programs that rely on self-assessment of these factors. There was no evidence that case and control participants differed in their level of agreement, which suggests that differential misclassification of phenotype characteristics is unlikely to have occurred in this case-control study, and thus minimizing the potential for risk estimates

to be biased. However, there was some evidence that women had better accuracy than men at reporting their own skin color, and that participants from lower latitudes had worse agreement for eye color and, to a lesser degree skin color, compared with participants from higher latitudes.

Our study found excellent agreement for eye color and moderate agreement for hair color when comparing self-report with dermatologist assessment. A study of 114 participants ages 40 to 69 years from a cohort in Queensland, Australia, reported moderate agreement for eye color ( $\kappa = 0.58$ ) and hair color ( $\kappa = 0.63$ ; ref. 25). In the United Kingdom, Melia and colleagues (26) also found moderate agreement for self-report versus dermatologist assessment of hair color.

There was fair to moderate agreement between self-reported skin color and continuous measures of spectrophotometer-derived



**Figure 2.** Plot comparing self-reported and clinically measured total nevus counts on the back. Intraclass correlation (ICC) = 0.36; 95% CI, 0.30–0.41.

**Table 4.** Intraclass correlations (ICC) comparing clinically measured whole-body nevus counts (2+ mm) with site-specific nevus counts

Body site	Males		Females	
	ICC	95% CI	ICC	95% CI
Posterior body <sup>a</sup>	0.87	0.84–0.89	0.84	0.82–0.86
Back	0.46	0.38–0.53	0.31	0.23–0.38
Both arms	0.49	0.42–0.56	0.57	0.52–0.62
Right arm only	0.26	0.17–0.35	0.32	0.24–0.39
Both thighs	0.33	0.24–0.42	0.40	0.33–0.47

<sup>a</sup>The classification of the different body sites is shown in Supplementary Fig. S1.

skin reflectance. However, when skin color reflectance values were categorized into very fair, fair, and olive/dark categories, the agreement was poor ( $\kappa = 0.12$ ). We used published cutoff points for ITA scores (19, 20) but a range of different reflectance values and cutoff points have been used in other studies (23, 27, 28) to classify skin color. Our study sample had a fairly homogenous skin type and a narrow range of skin reflectance values, which probably contributed to the poor agreement with self-reported skin color, as has been noted in other studies conducted among predominantly white populations (23). Studies with a greater diversity of skin color including non-white participants have demonstrated better correlation between self-report and reflectance spectrophotometry (15, 20).

Of the people who classified themselves as having olive or dark skin color in our study, only 21% had the same classification based on spectrophotometer assessment, whereas 54% were classified by spectrophotometry as having "fair" skin and 25% as having "very fair" skin. Similarly, a New Zealand study of 289 university students (20) found that 77% of participants who classified themselves as having medium skin color were in the fair skin color category when measured by spectrophotometry. In a convenience sample of 341 adults in Queensland, Australia, 47% of those categorized as fair skin color by spectrophotometry reported their own skin color as medium or olive, and agreement was summarized by a  $\kappa_w$  of 0.11 (95% CI, 0.01–0.21; ref. 23). Studies that have compared self-report with dermatologist (but not spectrophotometer) assessment of skin color have also found fair agreement (26, 29). However, contrary to our results of higher skin color agreement for women than men, Melia and colleagues reported better agreement for men (OR = 3.1; 95% CI, 1.4–6.8; ref. 26).

Number of nevi is one of the strongest risk factors for melanoma (8) but is often problematic for people to accurately report (30), partly due to confusion with solar lentigines, freckles, and seborrhoeic keratoses (11). Our study found fair to moderate agreement between self-reported and clinically assessed nevus counts. Participants underestimated their nevus counts on the back by more than half. Despite this underreporting, our data demonstrate that participants are able to adequately rank themselves; the correlation was 0.36 (95% CI, 0.30–0.41) comparing self-report and clinician counts on the back ( $\kappa_w$  0.28 when categorized into quartiles) and 0.43 (95% CI, 0.38–0.49) when participants classified themselves into descriptive (none, few, some, many) categories of whole-body nevus density. However, discrimination was modest between the "none" and "few" categories and thus consideration should be given to combining these categories in future risk analyses. Other studies that have used similar diagrams for classifying self-reports of whole-body nevus density have found comparable agreement with clinical assessment (25, 31). Interestingly, in the Queensland study (25), the

number of clinically measured nevi corresponding to each descriptive nevus category was less than half that recorded for the same categories in our study. This difference might be due to the younger age of participants in our study since nevi start to regress after early adulthood (32), differences in clinician reporting, or because our sample also included melanoma cases. Nevertheless, it highlights that caution should be used when combining data for self-reported nevus categories across studies, because although the data are suitable for ranking people according to their nevus count, they should not be used to estimate absolute number of nevi. In contrast with the underreporting of nevi in our study, a study of university students in Germany found that participants counted on average 14% more nevi on the arms than the examiners, although the discrepancy was larger for nonmedical students and those with fair skin (33).

Direct comparisons with previous validation studies are hampered by differences in study design, nevus-counting protocols, nevus definition and type, statistical methods, the sex and age of participants, and whether counts were conducted on the whole body or isolated to specific anatomical sites. Other validation studies of self-reported whole body nevus counts have generally reported poor (34) or fair agreement (29), although agreement tends to be higher when classified into fewer categories (35, 36). We did not find any differences in level of agreement for nevus counts by sex or case-control status, which is consistent with some studies (31, 33, 34); however, Lawson and colleagues (11) found that men were more likely than women to underreport large nevi.

For our secondary aim, we examined agreement between clinically measured nevus counts on the whole body to that from selected anatomical regions. This comparison is relevant for epidemiologic studies, which for reasons of time, cost, and feasibility, often use localized nevus count as a proxy measure of whole body nevus counts. We found excellent agreement between nevus counts on the entire back side of the body and total body nevus count. However, for studies where an extensive nevus count on the body is impractical, our results suggest that measuring nevus counts on both arms would be the best alternative, consistent with previous reports (37, 38). Nevus distribution is related to age, sex, and sun exposure; however, sex had little impact on anatomic site as a proxy measure for whole-body nevus count.

Strengths of our study include the large population-based sample of young adults, and comprehensive measures of pigmentation and nevus phenotype. We used spectrophotometer assessment of skin reflectance, which is free from observer bias and errors associated with recall. The clinical phenotype assessment was conducted by experienced dermatology trainees who were trained on the study protocol, and they used standardized clinical measurement tools including wig hair swatches, eye photographs, and a plastic size standard for counting nevi. However, clinician measurement still has an element of subjectivity and this may have influenced our results. In addition, we did not examine interexaminer reliability. In other studies, interexaminer reliability has been found to range from moderate to excellent between dermatologists (11, 39), but was poor when comparing experienced dermatologists with nonmedically trained interviewers (39). Another limitation was that the categories for measuring and classifying nevus and pigmentation phenotype did not always match perfectly between self-report and clinical assessment, and we had few self-reported nevus measures.

These results have implications for sun-related public health programs targeted at younger people. To improve the likely

success of future public health campaigns, it would be worthwhile to develop and test novel ways of improving self-assessment and simple, objective measurement of nevus and pigmentation phenotype. For example, technological advances would now enable the development of online self-assessment systems showing standardized photographs for comparison, or mobile phone apps that could detect skin color from real-time photographs and count freckles and nevi based on in-built algorithms. It is also of interest to determine to what extent people who underestimate their nevus and pigmentation phenotype also underestimate their susceptibility to melanoma, and how this correlates with sun-related behaviors. In the meantime, public health campaigns targeted at younger people should not rely too heavily on self-identification of high-risk status based on phenotypic characteristics.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Authors' Contributions

**Conception and design:** A.E. Cust, G.G. Giles, J.L. Hopper, M.A. Jenkins, G.J. Mann  
**Development of methodology:** A.E. Cust, J. Kelly, J.L. Hopper, G.J. Mann  
**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** A.E. Cust, H. Schmid, J. Kelly, J.F. Aitken, G.G. Giles, J.L. Hopper, G.J. Mann  
**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** A.E. Cust, K.M. Pickles, C. Goumas, T. Vu, J.L. Hopper, G.J. Mann  
**Writing, review, and/or revision of the manuscript:** A.E. Cust, K.M. Pickles, C. Goumas, T. Vu, E. Nagore, J. Kelly, J.F. Aitken, G.G. Giles, J.L. Hopper, M.A. Jenkins, G.J. Mann  
**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** A.E. Cust, K.M. Pickles, T. Vu, H. Schmid  
**Study supervision:** A.E. Cust, J. Kelly, G.J. Mann  
**Other (funding support):** A.E. Cust

### References

1. Australian Institute of Health and Welfare & Australasian Association of Cancer Registries, Cancer in Australia: an overview, 2012, in Cancer Series no. 74, Cat. no. CAN 70. AIHW: Canberra; 2012.
2. Homer MJ, Ries LAG, Krapcho M, Neyman N, Aminou R, Howlander N, et al. SEER Cancer Statistics Review, 1975–2006. 2009, National Cancer Institute: Bethesda, MD. Available from: [http://seer.cancer.gov/csr/1975\\_2006/](http://seer.cancer.gov/csr/1975_2006/).
3. Cancer Research UK. CancerStats - Cancer Statistics for the UK. [Accessed July 6, 2014]; Available from: <http://info.cancerresearchuk.org/cancerstats/>.
4. Australian Institute of Health and Welfare (AIHW), ACIM (Australian Cancer Incidence and Mortality) Books. Canberra: AIHW; 2010.
5. Armstrong BK, Kricger A. How much melanoma is caused by sun exposure? *Melanoma Res* 1993;3:395–401.
6. Green AC, Williams GM, Logan V, Stratton GM. Reduced melanoma after regular sunscreen use: randomized trial follow-up. *J Clin Oncol* 2011;29:257–63.
7. Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 2009;27:6199–206.
8. Gandini S, Sera F, Cattaruzza MS, Pasquini P, Abeni D, Boyle P, et al. Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. *Eur J Cancer* 2005;41:28–44.
9. Gandini S, Sera F, Cattaruzza MS, Pasquini P, Zanetti R, Masini C, et al. Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. *Eur J Cancer* 2005;41:2040–59.
10. Titus-Ernstoff L, Thorn M, Tosteson TD, Brahm EM, Yuen J, Baron JA, et al. The accuracy of skin self-examination for atypical nevi. *Epidemiology* 1996;7:619–23.
11. Lawson DD, Moore DH II, Schneider JS, Sagebiel RW. Nevus counting as a risk factor for melanoma: comparison of self-count with count by physician. *J Am Acad Dermatol* 1994;31:438–44.
12. Cust AE, Schmid H, Maskiell JA, Jetann J, Ferguson M, Holland EA, et al. Population-based, case-control-family design to investigate genetic and environmental influences on melanoma risk: Australian Melanoma Family Study. *Am J Epidemiol* 2009;170:1541–54.
13. English DR, MacLennan R, Rivers J, Kelly J, Armstrong BK. Epidemiological studies of melanocytic naevi: protocol for identifying and recording naevi. 1990, International Agency for Research on Cancer (IARC) internal report No. 90/002: Lyon (France).
14. Dwyer T, Blizzard L, Ashbolt R. Sunburn associated with increased number of nevi in darker as well as lighter skinned adolescents of northern European descent. *Cancer Epidemiol Biomarkers Prev* 1995;4:825–30.
15. Eilers S, Bach DQ, Gaber R, Blatt H, Guevara Y, Nitsche K, et al. Accuracy of self-report in assessing Fitzpatrick skin phototypes I through VI. *JAMA Dermatol* 2013;149:1289–94.
16. Weatherall IL, Coombs BD. Skin color measurements in terms of CIELAB color space values. *J Invest Dermatol* 1992;99:468–73.
17. Creech LL, Mayer JA. Ultraviolet radiation exposure in children: a review of measurement strategies. *Ann Behav Med* 1997;19:399–407.
18. Dodd AT, Morelli J, Mokrohisky ST, Asdigian N, Byers TE, Crane LA. Melanocytic nevi and sun exposure in a cohort of Colorado children: anatomic distribution and site-specific sunburn. *Cancer Epidemiol Biomarkers Prev* 2007;16:2136–43.
19. Del Bino S, Sok J, Bessac E, Bernerd F. Relationship between skin response to ultraviolet exposure and skin color type. *Pigment Cell Res* 2006;19:606–14.
20. Reeder AI, Hammond VA, Gray AR. Questionnaire items to assess skin color and erythral sensitivity: reliability, validity, and "the dark shift." *Cancer Epidemiol Biomarkers Prev* 2010;19:1167–73.
21. Cohen J. Weighted Kappa: nominal scale agreement with provision for scaled agreement or partial credit. *Psychol Bull* 1968;70:213–20.

### Acknowledgments

The authors thank all of the participants, the work and dedication of the research coordinators, interviewers, examiners, and data management staff, including Judith Maskiell, Jackie Arbuckle, Steven Columbus, Michaela Lang, Helen Rodais, Caroline Ellis (Centre for Epidemiology & Biostatistics, School of Population Health, University of Melbourne, Melbourne, Australia); Carol El Hayek, Lynne Morgan, Joanne Roland, Emma Tyler, Jodi Barton, Caroline Watts, and Lesley Porter (Westmead Institute of Cancer Research, University of Sydney at Westmead Millennium Institute, Sydney, Australia); Jodie Jetann, Megan Ferguson, Michelle Hillcoat, Kellie Holland, Pamela Saunders, Joan Roberts, and Sheree Tait (Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Spring Hill, Brisbane, Australia); Anil Kurien, Clare Patterson, Caroline Thoo, Sally de Zwaan, Angelo Sklavos, Shobhan Manoharan, Jenny Cahill, and Sarah Brennan (skin examiners). The authors also thank Bruce Armstrong and Richard Kefford for their role as Investigators on the Australian Melanoma Family Study.

### Grant Support

This work was supported by the National Health and Medical Research Council of Australia (NHMRC; project grants 566946, 107359, 211172 and program grant number 402761 to G.J. Mann); the Cancer Council New South Wales (project grant 77/00, 06/10), the Cancer Council Victoria, and the Cancer Council Queensland (project grant 371); and the US NIH (NIH RO1 grant CA-83115-01A2 and 2R01CA083115-11A1 to the International Melanoma Genetics Consortium, GenoMEL). A.E. Cust is the recipient of a NHMRC Career Development Fellowship (1063593) and a Cancer Institute NSW Early Career Development Fellowship (10/ECF/2-06).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 23, 2014; revised January 15, 2015; accepted January 17, 2015; published OnlineFirst January 27, 2015.

22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
23. Harrison SL, Buttner PG. Do all fair-skinned Caucasians consider themselves fair? *Prev Med* 1999;29:349–54.
24. Stanton WR, Janda M, Baade PD, Anderson P. Primary prevention of skin cancer: a review of sun protection in Australia and internationally. *Health Promot Int* 2004;19:369–78.
25. Morze CJ, Olsen CM, Perry SL, Jackman LM, Ranieri BA, O'Brien SM, et al. Good test-retest reproducibility for an instrument to capture self-reported melanoma risk factors. *J Clin Epidemiol* 2012;65:1329–36.
26. Melia J, Harland C, Moss S, Eiser JR, Pendry L. Feasibility of targeted early detection for melanoma: a population-based screening study. *Br J Cancer* 2000;82:1605–9.
27. Pershing LK, Tirumala VP, Nelson JL, Corlett JL, Lin AG, Meyer LJ, et al. Reflectance spectrophotometer: the dermatologists' sphygmomanometer for skin phototyping? *J Invest Dermatol* 2008;128:1633–40.
28. Dwyer T, Muller HK, Blizzard L, Ashbolt R, Phillips G. The use of spectrophotometry to estimate melanin density in Caucasians. *Cancer Epidemiol Biomarkers Prev* 1998;7:203–6.
29. Richtig E, Santigli E, Fink-Puches R, Weger W, Hofmann-Wellenhof R. Assessing melanoma risk factors: how closely do patients and doctors agree? *Public Health* 2008;122:1433–9.
30. Hamidi R, Peng D, Cockburn M. Efficacy of skin self-examination for the early detection of melanoma. *Int J Dermatol* 2010;49:126–34.
31. Walter SD, Marrett LD, Hertzman C. Reliability of interviewer and subject assessments of nevus counts in a study of melanoma. *J Clin Epidemiol* 1991;44:633–40.
32. MacKie RM, English J, Aitchison TC, Fitzsimons CP, Wilson P. The number and distribution of benign pigmented moles (melanocytic naevi) in a healthy British population. *Br J Dermatol* 1985;113:167–74.
33. Fiessler C, Pfahlberg A, Li J, Uter W, Gefeller O. Accuracy and reliability of naevus self-counts. *Melanoma Res* 2014;Jul 10. [Epub ahead of print].
34. Buettner PG, Garbe C. Agreement between self-assessment of melanocytic nevi by patients and dermatologic examination. *Am J Epidemiol* 2000;151:72–7.
35. Jackson A, Wilkinson C, Ranger M, Pill R, August P. Can primary prevention or selective screening for melanoma be more precisely targeted through general practice? A prospective study to validate a self administered risk score. *BMJ* 1998;316:34–8.
36. Carli P, De Giorgi V, Nardini P, Mannone F, Palli D, Giannotti B. Melanoma detection rate and concordance between self-skin examination and clinical evaluation in patients attending a pigmented lesion clinic in Italy. *Br J Dermatol* 2002;146:261–6.
37. English JS, Swerdlow AJ, Mackie RM, O'Doherty CJ, Hunter JA, Clark J, et al. Site-specific melanocytic naevus counts as predictors of whole body naevi. *Br J Dermatol* 1988;118:641–4.
38. Gallus S, Naldi L, Carli P, La Vecchia C. Nevus count on specific anatomic sites as a predictor of total body count: a survey of 3,406 children from Italy. *Am J Epidemiol* 2007;166:472–8.
39. Byles JE, Hennrikus D, Sanson-Fisher R, Hersey P. Reliability of naevus counts in identifying individuals at high risk of malignant melanoma. *Br J Dermatol* 1994;130:51–6.