



A Model to Predict the Risk of Keratinocyte Carcinomas

David C. Whiteman¹, Bridie S. Thompson¹, Aaron P. Thrift², Maria-Celia Hughes¹, Chiho Muranushi^{1,3}, Rachel E. Neale¹, Adele C. Green^{1,4} and Catherine M. Olsen¹, for the QSkin Study

Basal cell and squamous cell carcinomas of the skin are the commonest cancers in humans, yet no validated tools exist to estimate future risks of developing keratinocyte carcinomas. To develop a prediction tool, we used baseline data from a prospective cohort study ($n = 38,726$) in Queensland, Australia, and used data linkage to capture all surgically excised keratinocyte carcinomas arising within the cohort. Predictive factors were identified through stepwise logistic regression models. In secondary analyses, we derived separate models within strata of prior skin cancer history, age, and sex. The primary model included terms for 10 items. Factors with the strongest effects were >20 prior skin cancers excised (odds ratio 8.57, 95% confidence interval [95% CI] 6.73–10.91), >50 skin lesions destroyed (odds ratio 3.37, 95% CI 2.85–3.99), age ≥ 70 years (odds ratio 3.47, 95% CI 2.53–4.77), and fair skin color (odds ratio 1.75, 95% CI 1.42–2.15). Discrimination in the validation dataset was high (area under the receiver operator characteristic curve 0.80, 95% CI 0.79–0.81) and the model appeared well calibrated. Among those reporting no prior history of skin cancer, a similar model with 10 factors predicted keratinocyte carcinoma events with reasonable discrimination (area under the receiver operator characteristic curve 0.72, 95% CI 0.70–0.75). Algorithms using self-reported patient data have high accuracy for predicting risks of keratinocyte carcinomas.

Journal of Investigative Dermatology (2016) **136**, 1247–1254; doi:10.1016/j.jid.2016.02.008

INTRODUCTION

Keratinocyte carcinomas (KCs) (specifically, basal cell carcinomas (BCC) and squamous cell carcinomas (SCC) of the skin) are the most common cancers in humans. Although often regarded as unimportant cancers, KCs cause considerable morbidity. Each year, 1.9% of the US adult population is estimated to receive treatment for KC, rising to 6.9% of the population aged >65 years (U.S. Department of Health and Human Services, 2014). As such, KCs are the most costly cancers in the US Medicare population, contributing to the estimated total direct treatment costs of \$4.3 billion each year (Housman et al., 2003; U.S. Department of Health and Human Services, 2014). Globally, the highest rates of BCC ($>1000 \times 10^{-5}$ person-years) and SCC (387×10^{-5} person-years) are observed in Australia, although KC incidence rates exceed 100×10^{-5} person-years in most fair-skinned populations around the world (Lomas et al., 2012).

The enormous burden of skin cancers underscores the need to find better ways to control these conditions, both through

primary prevention and early detection (U.S. Department of Health and Human Services, 2014). Reducing hazardous exposure to solar ultraviolet radiation is accepted as the mainstay of primary prevention in the general population, but there is less certainty about how to deploy medical services for skin cancer control. At one end of the early detection spectrum is formal population-based screening. To date, only Germany has embarked upon a nationwide program of screening for skin cancer (Choudhury et al., 2012), although evidence is emerging that the benefits may not justify the costs (Katalinic et al., 2015; Stang and Jockel, 2015). In other jurisdictions, guidelines recommend that patients at high risk for skin cancer undergo periodic surveillance, with the implied but unstated advice that those at low risk receive usual care (Cancer Council Australia, 2007; U.S. Preventive Services Task Force, 2009). With the advent of promising chemoprevention strategies for skin cancer (Chen et al., 2015), reliable risk stratification tools are needed to identify patients most likely to benefit from clinical intervention. Presently, clinicians must rely on their own experience to estimate a patient's future risk of skin cancer, because unlike for melanoma (Olsen et al., 2015), no prediction tools for these cancers have been developed or validated. Thus, we aimed to develop and validate a risk prediction model for quantifying the probability of being treated for a KC, regardless of histological subtype. Our approach was to use self-reported information of the type that can be elicited remotely, so that patients might be assessed and triaged before consulting a physician.

RESULTS

The eligible cohort comprised 38,726 participants with no prior history of melanoma, of whom 56% were women and the mean age was 56.2 years (standard deviation 8.1). Most

¹Population Health Department, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia; ²Department of Medicine and Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, USA; ³School of Population Health, The University of Queensland, Herston, Queensland, Australia; and ⁴Cancer Research UK Manchester Institute and Institute of Inflammation and Repair, University of Manchester, Manchester, UK

Correspondence: David C. Whiteman, Cancer Control Group, QIMR Berghofer Medical Research Institute, Locked Bag 2000, RBWH, Queensland 4029, Australia. E-mail: david.whiteman@qimrberghofer.edu.au

Abbreviations: AUROC, area under the receiver operator characteristic curve; BCC, basal cell carcinomas; KC, keratinocyte carcinoma; SCC, squamous cell carcinomas

Received 16 November 2015; revised 19 January 2016; accepted 1 February 2016; accepted manuscript published online 22 February 2016; corrected proof published online 22 March 2016

participants reported white European ancestry (n = 34,579, 93%) and most were born in Australia (30,054, 81%). Median follow-up was 36 months (min 20; max 44). Distributions of key factors in the derivation and validation datasets are presented in Table 1.

There were 4,237 (16%) participants in the derivation dataset with at least one surgical excision for a confirmed BCC or SCC during the follow-up period. In univariate analyses, 23 items were significantly associated with the occurrence of KC (Supplementary Table S1 online). After conducting backward stepwise regression within the imputed derivation datasets (Supplementary Table S2 online), and then testing whether adding further candidate terms significantly improved the fit of the model, we derived a final model with terms for 10 predictors (age, sex, smoking status, ethnicity, skin color, tanning ability, freckling tendency, number of sunburns <10 years, number of previous skin cancers excised, and number of previous skin lesions destroyed; Table 2). None of the pairwise interaction terms was statistically significant at the *P* < 0.05 level, nor did they

substantially modify the Akaike Information Criterion or the area under the receiver operator characteristic curve (AUROC); hence they were not retained in the final models. Strongest effects were observed for >20 prior skin cancers excised, >50 skin lesions destroyed, age >70 years, and fair skin color. Discrimination in the validation dataset was high (AUROC 0.80, 95% confidence interval 0.79–0.81; Figure 1). Although the model appeared well calibrated overall, there was some evidence that true risks were underestimated at the low end of the scale, and that risks were marginally overestimated in the highest categories (Figure 2).

Cumulative incidence plots demonstrated that the risk of KC events was strongly predicted by the self-reported history of prior skin cancer excisions (Figure 3). Thus, we derived separate models according to the absence or presence of a past skin cancer history. Among those in the derivation dataset with no prior history of skin cancer excisions (n = 16,021), we derived a predictive model with good discrimination (AUROC 0.72, 95% confidence interval 0.70–0.75;

Table 1. Characteristics of study participants in derivation and validation datasets

Characteristic	Derivation dataset			Validation dataset		
	Controls (N = 21,605)	KC cases (N = 4,237)	Total (N = 25,842)	Controls (N = 10,773)	KC cases (N = 2,111)	Total (N = 12,884)
Age group						
40–49	6,309 (29.2)	616 (15.4)	6,925 (26.8)	3,136 (29.1)	273 (12.9)	3,409 (26.5)
50–59	8,371 (38.8)	1,471 (34.7)	9,842 (38.1)	4,183 (38.8)	722 (34.2)	4,905 (38.1)
60–69	6,784 (31.4)	2,063 (48.7)	8,847 (34.2)	3,377 (31.4)	1,069 (50.6)	4,446 (34.5)
70+	141 (0.7)	87 (2.1)	228 (0.9)	77 (0.7)	47 (2.2)	124 (1.0)
Sex						
Female	12,327 (57.1)	1,933 (45.6)	14,260 (55.2)	6,094 (56.6)	975 (46.2)	7,069 (54.9)
Male	9,278 (42.9)	2,304 (53.7)	11,582 (44.8)	4,679 (43.4)	1,136 (53.8)	5,815 (45.1)
Ethnicity						
Non-white	1,500 (6.9)	85 (2.0)	1,585 (6.9)	747 (6.9)	42 (2.0)	789 (6.1)
White	19,889 (92.1)	4,115 (97.1)	24,004 (92.9)	9,917 (92.1)	2,055 (97.4)	11,972 (92.9)
Missing	216 (1.0)	37 (0.9)	253 (1.0)	109 (1.0)	14 (0.7)	123 (1.0)
Born in Australia						
No	4,537 (21.0)	592 (14.0)	5,129 (21.0)	2,272 (21.1)	280 (13.3)	2,552 (19.8)
Yes	17,062 (78.0)	3,641 (85.9)	20,703 (78.0)	8,495 (78.9)	1,831 (86.7)	10,326 (80.2)
Missing	6 (<0.1)	4 (0.1)	10 (<0.1)	6 (0.1)	0 (0)	6 (0.1)
Private health insurance						
No	7,311 (33.8)	1,313 (31.0)	8,624 (33.4)	3,564 (33.1)	624 (29.6)	4,188 (32.5)
Yes	14,191 (65.7)	2,900 (68.4)	17,091 (66.1)	7,168 (66.5)	1,480 (70.1)	8,648 (67.1)
Missing	103 (0.5)	24 (0.6)	127 (0.5)	41 (0.4)	7 (0.3)	48 (0.4)
Education						
University degree	5,341 (24.7)	855 (20.2)	6,196 (25.0)	2,691 (25.0)	410 (19.4)	3,101 (24.1)
Certificate or diploma	4,070 (18.8)	691 (16.3)	4,761 (18.4)	2,145 (19.9)	328 (15.5)	2,473 (19.2)
Trade	1,954 (9.0)	414 (9.8)	2,368 (9.2)	963 (8.9)	213 (10.1)	1,176 (17.8)
Secondary school	3,974 (18.4)	736 (17.4)	4,710 (18.2)	1,906 (17.7)	386 (18.3)	2,292 (17.8)
Intermediate school	3,200 (14.8)	764 (18.0)	3,964 (15.3)	1,550 (14.4)	398 (18.9)	1,948 (15.1)
Primary school	1,575 (7.3)	417 (9.8)	1,992 (7.7)	785 (7.3)	194 (9.2)	979 (7.6)
Missing	1,491 (6.9)	360 (8.5)	1,851 (7.2)	733 (6.8)	182 (8.6)	915 (7.1)
Smoking status						
Never	11,763 (54.5)	2,260 (53.3)	14,023 (54.3)	5,944 (55.2)	1,155 (54.7)	7,099 (55.1)
Past	7,610 (35.2)	1,523 (36.0)	9,133 (35.3)	3,760 (34.9)	743 (35.2)	4,503 (35.0)
Current	2,120 (9.8)	424 (10.0)	2,544 (9.8)	1,002 (9.3)	203 (9.6)	1,205 (9.4)
Missing	112 (0.5)	30 (0.7)	142 (0.5)	67 (0.6)	67 (0.6)	77 (0.6)

Abbreviation: KC, keratinocyte carcinoma.

Table 2. Specification and performance of models to predict risk of keratinocyte carcinoma within 3 years: full derivation dataset (N = 25,842)

Risk factor	OR (95% CI) ¹	P-value ²
Age group		
40–49	1.00 (Ref)	
50–59	1.25 (1.13, 1.40)	
60–69	1.76 (1.57, 1.96)	
70+	3.47 (2.53, 4.77)	<0.0001
Sex		
Female	1.00 (Ref)	
Male	1.31 (1.21, 1.41)	<0.0001
Ethnicity		
Non-white	1.00 (Ref)	
White	1.72 (1.36, 2.17)	<0.0001
Skin color		
Dark	1.00 (Ref)	
Medium	1.50 (1.22, 1.84)	
Fair	1.75 (1.42, 2.15)	<0.0001
Tanning ability		
Deeply tan	1.00 (Ref)	
Moderately tan	1.20 (1.05, 1.35)	
Lightly tan	1.17 (1.05, 1.30)	
Not tan	1.16 (0.98, 1.37)	0.0186
Freckling tendency		
None	1.00 (Ref)	
A few	1.00 (0.91, 1.09)	
Some	1.06 (0.94, 1.18)	
Many	1.21 (1.04, 1.41)	0.0557
Number of sunburns <10 years		
Never	1.00 (Ref)	
1–5	1.03 (0.93, 1.14)	
6–10	1.07 (0.94, 1.21)	
11–20	1.11 (0.96, 1.29)	
21–50	0.79 (0.65, 0.97)	
>50	0.88 (0.65, 1.18)	0.0148
Number of previous skin cancers excised		
None	1.00 (Ref)	
1	1.79 (1.60, 2.01)	
2–10	3.32 (3.02, 3.66)	
11–20	7.74 (6.38, 9.38)	
>20	8.57 (6.73, 10.91)	<0.0001
Number of previous skin lesions destroyed		
None	1.00 (Ref)	
1–5	1.75 (1.57, 1.94)	
6–10	2.35 (2.06, 2.68)	
11–20	2.81 (2.44, 3.23)	
21–50	2.72 (2.33, 3.17)	
>50	3.37 (2.85, 3.99)	<0.0001
Smoking status		
Never	1.00 (Ref)	
Former	0.95 (0.87, 1.03)	
Current	1.17 (1.03, 1.33)	0.0052
AUROC (95% CI) ³		
Derivation	0.79 (0.78, 0.80)	
Validation	0.80 (0.79, 0.81)	

¹Odds ratios (and 95% confidence intervals) were adjusted for all other terms in the table. Missing values were imputed as per protocol.

²Type III *P*-value from first imputation dataset (see [Supplementary Table S2](#) online).

³AUROC, area under the receiver operating characteristic curve; CI, confidence interval.

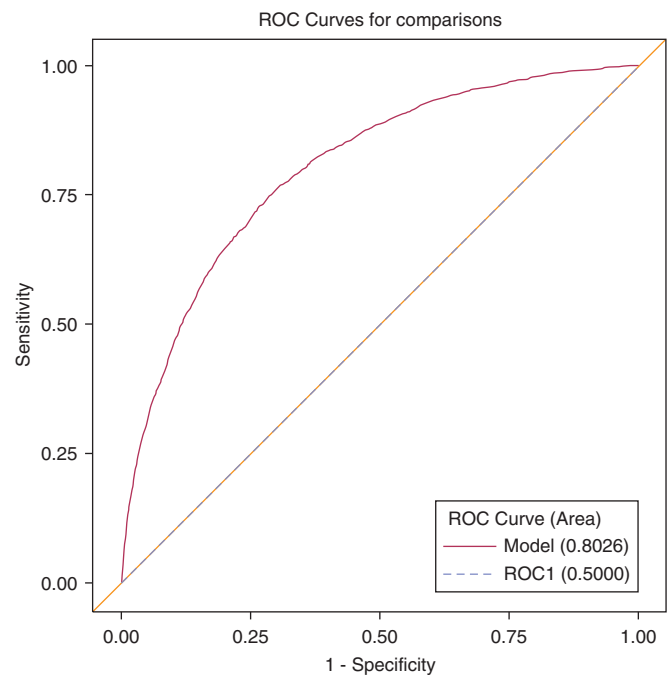


Figure 1. Discrimination performance of the keratinocyte carcinoma prediction model in the validation cohort. Performance is assessed by the area-under-the-receiver-operating-characteristic curve (AUROC). The x-axis denotes the false-positive rate and the y-axis denotes the true-positive rate. The diagonal line represents the “line of non-discrimination” which is equivalent to the prediction tool performing no better than random chance.

[Supplementary Figures S1 and S2](#) online) that comprised 10 statistically significant terms, namely age, sex, ethnicity, skin color, tanning ability, freckling, sunburns < 10 years, sunburns > 20 years, number of previous skin lesions destroyed, and family history of melanoma ([Table 3](#)). Among those with a self-reported history of at least one prior skin cancer excision ($n = 9,650$), the most parsimonious prediction model comprised eight terms (namely number of previous skin cancers excised, number of previous skin lesions destroyed, age, sex, smoking status, skin color, freckling tendency, and sunburns < 10 years) ([Table 3](#)). Discrimination in the validation cohort was 0.72 (95% confidence interval 0.70–0.73) ([Supplementary Figures S3 and S4](#) online).

In supplementary analyses, we derived prediction models within strata of age and sex ([Supplementary Table S3](#) online). Discrimination of the models in the validation datasets was similarly high across all models (AUROC approximately 0.80), although the terms included in the prediction models differed. Models restricted to older people (≥ 60 years) and women had fewer terms (seven each) than models restricted to younger people (10 terms) and men (11 terms). Across all four strata, terms were retained for age, ethnicity, skin color, numbers of skin cancers excised, and numbers of skin lesions destroyed.

Finally, to assess the utility of using the prediction model to guide clinical management, we calculated the proportion of patients in the source population that would be considered for clinical action (e.g., chemoprevention, surveillance program, etc.) at different risk score thresholds as predicted by the model ([Table 4](#)). Assuming a stringent scenario in which clinical action would be taken only for patients with a very

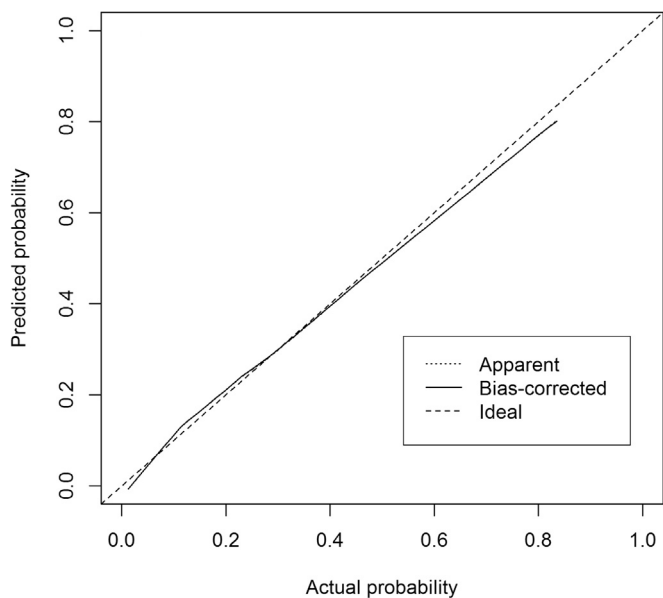


Figure 2. Calibration performance of the keratinocyte carcinoma prediction model in the validation cohort. The x-axis denotes the observed probability and the y-axis denotes the predicted probability. The plot presents the bootstrap-corrected calibration curves averaged over 500 replications.

high risk of developing skin cancer (e.g., 3-year risk score ≥ 0.7), then only approximately 1.3% of patients would be affected. Such a regimen would have extremely high specificity (99.5%) but very low sensitivity (5.4%), missing almost 95% of future cases of skin cancer. In contrast, selecting patients with 3-year risk score threshold ≥ 0.2 would permit 72.5% of patients to avoid clinical action, while yielding 64.5% of future cases with reasonable specificity (79.7%). The Youden Index was optimized ($J = 0.46$) at a predicted 3-year risk score of 0.13.

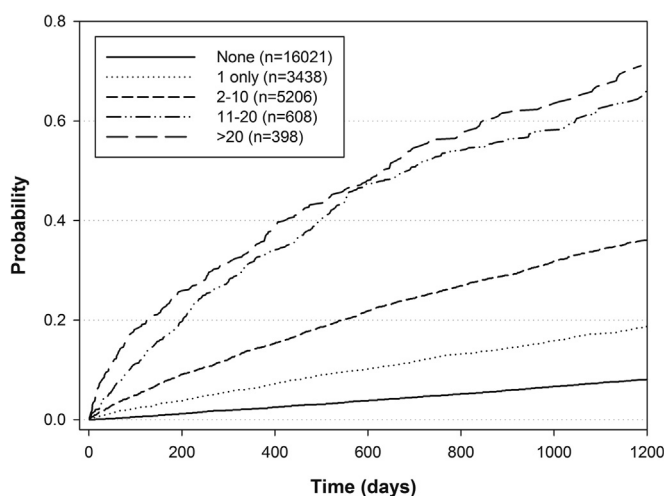


Figure 3. Cumulative incidence of keratinocyte carcinoma by category of the self-reported number of excised skin cancers before enrolment (171 participants had missing data for this item). The x-axis denotes time in study from date of consent and the y-axis denotes cumulative incidence (risk) of first excision of keratinocyte cancer. Separate cumulative incidence curves are presented for each of five groups of participants according to their self-reported prior history of skin cancer.

DISCUSSION

We used data from a large cohort in the high incidence setting of Queensland, Australia, to develop a tool for predicting the 3-year risk of developing KC. We found that information on 10 items yielded a risk prediction index for KCs with very high discrimination (AUROC 0.79). This level of discrimination is among the highest reported in a validation dataset for a cancer prediction index. Although statistical techniques are not yet available for testing the calibration performance of prediction tools in very large datasets (Paul et al., 2013), visual inspection of the plots suggests that the predicted probabilities accord favorably with the observed risk of KC events. Overall, the strongest predictors of future risk were a past history of an excised skin cancer or a destroyed actinic skin lesion, and age. Smoking status and a number of pigmentary factors including skin color, tanning ability, and freckling tendency also independently and significantly improved the fit of the model. In separate analyses restricted to participants with no self-reported skin cancer history, the predictive factors were largely the same, although smoking status was dropped from that model and three other factors (sunburns in childhood, sunburns in adulthood, and family history of melanoma) were retained. Interestingly, apart from sunburns, measures of past or current sun exposure were not retained in any models, despite their strong associations with KC events in univariate analyses. Several nonexclusive explanations are likely. First, a past history of skin cancer is essentially a proxy measure for high-level cumulative sun exposure. Secondly, in populations exposed to high levels of ambient solar ultraviolet radiation (such as Australia or southern USA), the role of pigmentation factors risk is likely accentuated because these factors strongly determine the actual dose of ultraviolet radiation received at the level of the target cell. Thirdly, because phenotypic measures are reported with higher reproducibility than sun exposure (both in this sample (Morze et al., 2012) and others (English et al., 1998; van der Mei et al., 2006; Yu et al., 2009)), they are less prone to misclassification and thus, on average, more likely to discriminate cases from controls.

The models described here are conceived to be clinical aids for stratifying patients according to their probability of developing skin cancer requiring surgical excision. As such, we used self-reported data of the type that can be collected remotely before consultation, and we did not discriminate between the development of BCC or SCC because initial management is similar for both. In addition to the pragmatic design, other strengths include the very large sample size (allowing separate datasets for deriving and validating models) and the use of a validated survey instrument for collecting baseline data that was specifically intended to assess skin cancer risk factors. We used health administration data ensuring virtually complete follow-up for all such events within the cohort. Moreover, we have shown that the eight items for excision of histologically confirmed KC have very high concordance (approximately 97%) with histopathologic diagnoses obtained independently from pathology records (Thompson et al., 2015). We consider it unlikely that any controls would be misclassified as cases, although it is possible that some study participants who did develop invasive KCs during follow-up may have been misclassified as

Table 3. Specification and performance of models to predict risk of keratinocyte carcinoma within 3 years, by strata of self-reported prior skin cancer history (N = 25,671)¹

Risk factor	Prior history of skin cancer excision (self-report)			
	Yes (N = 9,650)		No (N = 16,021)	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Age group				
40–49	1.00 (Ref)		1.00 (Ref)	
50–59	1.08 (0.93, 1.25)		1.53 (1.30, 1.81)	
60–69	1.47 (1.27, 1.69)		2.26 (1.90, 2.69)	
70+	3.10 (2.08, 4.62)	<0.0001	4.42 (2.55, 7.64)	<0.0001
Sex				
Female	1.00 (Ref)		1.00 (Ref)	
Male	1.26 (1.14, 1.38)	<0.0001	1.39 (1.22, 1.57)	<0.0001
Ethnicity				
Non-white	–		1.00 (Ref)	
White	–	n/s ²	2.07 (1.44, 2.96)	<0.0001
Skin color				
Dark	1.00 (Ref)		1.00 (Ref)	
Medium	1.60 (1.18, 2.16)		1.32 (0.99, 1.76)	
Fair	1.99 (1.48, 2.68)	<0.0001	1.44 (1.08, 1.92)	0.0405
Tanning ability				
Deeply tan	–		1.00 (Ref)	
Moderately tan	–		1.31 (1.10, 1.55)	
Lightly tan	–		1.34 (1.08, 1.65)	
Not tan	–	n/s	1.34 (0.98, 1.84)	0.0122
Freckling tendency				
None	1.00 (Ref)		1.00 (Ref)	
A few	0.89 (0.80, 1.00)		1.20 (1.04, 1.39)	
Some	0.94 (0.82, 1.07)		1.31 (1.08, 1.58)	
Many	1.10 (0.93, 1.32)	0.0543	1.41 (1.06, 1.86)	0.0094
Number of sunburns aged <10 years				
Never	1.00 (Ref)		1.00 (Ref)	
1–5	0.94 (0.82, 1.08)		1.09 (0.92, 1.29)	
6–10	0.95 (0.81, 1.12)		1.20 (0.98, 1.48)	
11–20	0.92 (0.76, 1.11)		1.37 (1.07, 1.75)	
21–50	0.63 (0.49, 0.80)		1.18 (0.81, 1.71)	
>50	0.74 (0.53, 1.05)	0.0029	0.82 (0.45, 1.51)	0.0972
Number of sunburns aged >20 years				
Never	–		1.00 (Ref)	
1–5	–		1.15 (0.96, 1.37)	
6–10	–		1.11 (0.89, 1.39)	
11–20	–		1.31 (1.01, 1.71)	
21–50	–		1.49 (1.04, 2.14)	
>50	–	n/s	0.74 (0.36, 1.52)	0.089
Number of previous skin cancers excised				
None	–		–	
1	1.00 (Ref)		–	
2–10	1.94 (1.74, 2.17)		–	
11–20	4.69 (3.84, 5.74)		–	
>20	5.14 (4.01, 6.60)	<0.0001	–	

(continued)

Table 3. Continued

Risk factor	Prior history of skin cancer excision (self-report)			
	Yes (N = 9,650)		No (N = 16,021)	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Number of previous skin lesions destroyed				
None	1.00 (Ref)		1.00 (Ref)	
1–5	1.45 (1.22, 1.71)		1.79 (1.55, 2.06)	
6–10	1.97 (1.65, 2.37)		2.27 (1.82, 2.84)	
11–20	2.25 (1.87, 2.71)		3.30 (2.58, 4.21)	
21–50	2.30 (1.89, 2.79)		3.18 (2.32, 4.36)	
>50	2.89 (2.35, 3.54)	<0.0001	4.36 (2.81, 6.76)	<0.0001
Family history of melanoma				
No	–		1.00 (Ref)	
Yes	–	n/s	1.19 (1.03, 1.38)	0.0126
Smoking status				
Never	1.00 (Ref)		–	
Former	0.96 (0.87, 1.07)		–	
Current	1.21 (1.03, 1.43)	0.0244	–	n/s
AUROC (95% CI)				
Derivation	0.71 (0.70, 0.72)		0.70 (0.68, 0.71)	
Validation	0.72 (0.70, 0.73)		0.72 (0.70, 0.75)	

Abbreviations: AUROC, area under the receiver operating characteristic curve; CI, confidence interval.

¹Missing values were imputed as per protocol. Participants with missing data for stratification variable were excluded from this analysis (n = 171).²Not statistically significant: term was not retained in the final risk prediction model.

controls in these analyses. This could happen if KCs were treated destructively or were removed by simple biopsies that did not attract one of the eight billing codes reserved for excisions of histologically confirmed KCs. The magnitude of bias of this type is difficult to estimate, although we note that approximately 15% of controls had at least one skin biopsy during follow-up. Such bias is most likely conservative because it would tend to diminish the differences between cases and controls, and thereby reduce the discrimination of the model. A potential limitation of the present analysis was the absence of histology data with which to separately predict risks of BCC and SCC. We aim to explore histology-specific prediction tools when such data become available.

Although population screening for melanoma and KCs was introduced in Germany in 2008, recent reports suggest that the program has not delivered the anticipated benefits (Katalinic et al., 2015; Stang and Jockel, 2015). Based on these findings, and the limited likelihood that any randomized trials for skin cancer screening will be conducted in the foreseeable future, there is growing interest in developing prediction tools to identify patients who may benefit from clinical interventions such as chemoprevention (Chen et al., 2015) or surveillance (Moons et al., 2009). An additional benefit of such tools, especially in high incidence populations or in settings where access to specialist care may be rationed, is to be able to triage people at low risk of disease back to their routine care providers. Prediction models for cancers of the breast (Rockhill et al., 2001), colon

Table 4. Sensitivity and specificity of the risk prediction model and proportion of “actionable” patients at incremental thresholds of risk scores for keratinocyte carcinoma within 3 years

Risk score threshold ¹	Sensitivity ² (%)	Specificity ³ (%)	“Actionable” proportion ⁴ (%)
0	100	0	100
≥0.1	82.7	60.7	64.3
≥0.2	64.5	79.7	27.5
≥0.3	49.6	88.2	17.9
≥0.4	34.5	93.9	10.7
≥0.5	20.9	97.3	5.6
≥0.6	13.3	98.6	3.3
≥0.7	5.4	99.5	1.3
≥0.8	0.3	99.9	0.1
≥0.9	0.0	100.0	0

¹Lower bound of the risk score for having a keratinocyte carcinoma event within 3 years as predicted by the full model.

²Proportion of cases who were correctly predicted to be a “case” (defined as having a risk score exceeding this threshold).

³Proportion of controls who were correctly predicted to be a “control” (defined as having a risk score below this threshold).

⁴Proportion of all participants (“the population”) who would be considered actionable at this level of risk score.

(Steffen et al., 2014), prostate (Eastham et al., 1999), and melanoma (Vuong et al., 2014; Whiteman, 2014) have been developed, but we are not aware of any published models that predict risk of KCs in the general population.

We therefore developed a prototype web-based application (see www.qskin.qimrberghofer.edu.au) to calculate a personal risk score. From a drop-down menu, a person selects the most appropriate response for each of the 10 items that were found to significantly predict risk of KC. The algorithm sums the beta-coefficients of the selected response items to generate a risk score, and then determines where that person’s score lies relative to the distribution of all risk scores in the QSkin cohort. For clinical utility, and to avoid perceptions of spurious precision, the tool reports a risk category rather than the actual score, as follows: (i) very much below average risk (bottom 20% of the risk distribution), (ii) below average risk (21st to 40th percentile of the risk distribution), (iii) about average risk (41st to 60th percentile of the risk distribution), (iv) above average risk (61st to 80th percentile of the risk distribution), (v) very much above average risk (top 20% of the risk distribution). Although the subsequent management of each patient will depend on his or her own particular circumstances, we believe this tool will aid clinicians and their patients in quantifying risk and deciding on an appropriate course of action. Although the external validity of this prediction tool remains to be determined, especially in settings where the incidence of KCs is lower than in Australia, it is notable that the factors conferring the greatest contribution to the risk score were not specific to this particular population. Moreover, the items to measure the factors were scaled across the full ranges of exposure, and are readily elicited by self-report either in a clinical encounter or remotely. In our future research, we aim to assess the impact of applying these prediction tools in diverse clinical settings.

MATERIAL AND METHODS

Study populations

The QSkin Study comprises a cohort of 43,794 men and women aged 40–69 years randomly sampled from the population of Queensland, Australia, in 2011 (overall participation fraction 23%). Full details of recruitment and baseline characteristics of the cohort have been published previously (Olsen et al., 2012). The Human Research Ethics Committee at the QIMR Berghofer Medical Research Institute approved the study, and all participants gave their informed and written consent to take part.

Data collection

At baseline, participants were asked to self-report information about demographic items, ethnicity, general medical history, pigmentary characteristics, history of sun exposure, use of tanning beds and past history of treatments for skin cancer, and other skin lesions. The characteristics of the questionnaire have been published, including measures of repeatability for study survey items (Morze et al., 2012).

The primary outcome for this analysis was excision of a histologically confirmed BCC or SCC (“KC event”). To identify KC events in the cohort, we obtained administrative claims data from Medicare Australia for all medical services provided to all participants who gave express written consent for data linkage (n = 40,383) between the date of consent (from November 2010) and censor date (30 June 2014). Medicare is Australia’s universal health care system that subsidizes virtually all medical services outside of the public hospital system for citizens and permanent residents, regardless of age, private health insurance status, or other factors (Olsen et al., 2014). Deterministic linkage to Medicare was conducted using Medicare number, name, address, and date of birth as identifiers. We defined KC cases as participants who received a medical service for one of eight item numbers (31255, 31260, 31265, 31270, 31275, 31280, 31285, 31290) for first surgical excision of BCC or SCC, for which the diagnosis must be confirmed by histology before the claim is lodged. We excluded 1,657 participants with confirmed pre-enrolment melanoma identified through record linkage to the population-based Queensland Cancer Registry; thus the final dataset for analysis comprised 38,726 participants.

Candidate predictor variables

Candidate predictor variables were selected a priori from the literature and practitioner input and included terms for demographic characteristics including age at enrollment, sex, place of birth, ethnicity, and place of residence as a child. We tested numerous self-reported measures of phenotype (including unexposed skin color, skin burning tendency, skin tanning tendency, eye color, hair color, categorical freckling density on the face at age 21, and categorical nevus burden at age 21), self-reported measures of sun exposure (including number of sunburns at ages <10 years, 10–20 years, and >20 years, numbers of hours spent outdoors on weekdays and weekend days in the past year), and frequency of use of tanning beds. Terms related to prior medical history included self-reported numbers of skin cancers excised surgically, self-reported numbers of skin lesions treated destructively, history of melanoma in close blood relatives, and frequency of aspirin use in the past year (see [Supplementary Table S1](#)).

Imputation of missing data

Missing values for most candidate predictor items occurred at prevalence <1% and was highest for educational attainment (7%). To avoid potential bias, we imputed missing values for candidate

predictors using the fully conditional specification method in PROC MI in SAS v9.4 (SAS Institute, Cary, NC), under the assumption that data were missing at random. We included all predictor variables and the outcome variable in the imputation step, specifying logistic regression to impute ordinal variables and linear regression for continuous variables. Imputation was run over five imputation cycles to generate five imputed datasets. We then performed the backward stepwise regression analyses on the five imputed datasets and combined the regression coefficients of the retained predictor variables using a modified version of "Rubin's rules" suitable for categorical variables (Sterne et al., 2009).

Statistical analysis

Our overall approach was to derive risk prediction models in a randomly selected sample of the dataset (two-thirds subset $n = 25,842$; hereafter "derivation sample"), and then to test the performance of models in the remaining sample ("validation sample," $n = 12,884$; see Figure 1). We compared KC cases and controls using χ^2 tests for categorical variables.

Risk model development. We used logistic regression for analysis, because the pragmatic consideration in the clinical setting is to determine the probability that a person with a given set of characteristics will require a KC excision within the next 3 years (i.e., become a "case"), or not. Although the logistic regression approach does not accommodate an analysis of competing risks of death, this is very unlikely to bias the estimates in our models because there were very few deaths during follow-up ($>98.5\%$ survival at 3 years). The very low rate of deaths was similar among cases and controls in both the development dataset and the validation dataset (development dataset: 1.49% vs. 1.29% in cases and controls, respectively, $P = 0.30$; validation dataset: 0.99% vs. 1.08% in cases and controls, respectively, $P = 0.74$). In the first phase, we included in the multivariate model those variables that were statistically significantly associated with KC at the 0.05 level in univariate analyses. We performed backward stepwise regression whereby those factors losing their significance at the 0.1 level in the multivariate analysis were dropped. In the second phase, those factors not significant in the univariate analyses were subsequently fitted to the multivariate model to identify any effects detectable only after adjusting for major risk factors.

Our analysis protocol stipulated that we were to perform analyses stratified by prior skin cancer history (yes vs. no); sex and age (<60 years vs. $60+$ years) on the a priori assumption that these factors would likely interact with predictor variables. Because of the substantial computational resources required to test all possible pairwise combinations of predictive factors, as well as the implications for multiple testing and the very low likelihood of meaningful gains in predictive value, we tested only for potential interactions between terms considered plausible and clinically relevant (age*sex; age*sunburns as a child; age*number of skin cancers excised; age*number of skin lesions destroyed; skin color*sunburns as a child; skin color*number of skin lesions destroyed). We included each pairwise interaction term in the primary analysis models and re-ran the logistic regression in the development dataset.

Model validation. We assessed the performance of derived prediction models using tests for discrimination and calibration. We evaluated discrimination using the AUROC (also known as the c-statistic) and its 95% confidence interval. The AUROC can be interpreted as the probability that the model will assign a higher

probability of developing KC to a randomly chosen participant who developed KC than to a randomly chosen participant who did not develop KC during follow-up. An AUROC of 0.5 indicates that the model discriminates no better than chance, whereas an AUROC of 1.0 indicates a perfectly discriminating model. We also assessed calibration that compares the observed proportions of KC versus controls within equally sized groups categorized according to their predicted probability from the model. When the average predicted risk within specified categories matches the proportion observed, the model is well calibrated. We plotted bootstrap-corrected calibration curves (averaged over 500 replications) to illustrate the model's fit across the range of predicted risk for KC compared with the observed outcome. We did not use the Hosmer-Lemeshow goodness-of-fit statistic because this measure is known to be overly sensitive with large samples sizes and as yet, no correction algorithms have been validated for samples $>25,000$ (Paul et al., 2013). We calculated the Youden index (J) for all data points in the sample to identify the "optimal" cutoff point at which both sensitivity and specificity are maximized.

In secondary analyses, we sought to establish whether predictive factors differed according to the past history of excision of skin cancer (self-reported), sex, and age (<60 years, ≥ 60 years), and so we repeated the process within strata of these variables. We excluded from the analyses stratified by self-reported prior history of skin cancer those participants with missing data for that item ($n = 171$, 0.6%). We used multiple imputation to complete missing values for all other variables.

Statistical computations were performed using SAS software (version 9.4; SAS Institute), and all tests for statistical significance were two-sided at $\alpha = 0.05$. Calibration was conducted using the rms package in R v3.0.1.

NOTE

An online version of the KC risk prediction calculator will be made available on our study website (www.qskin.qimrberghofer.edu.au) after publication of the article.

ORCID

David C. Whiteman: <http://orcid.org/0000-0003-2563-9559>

CONFLICT OF INTEREST

The authors state no conflicts of interest.

ACKNOWLEDGMENTS

We acknowledge the assistance of Medicare Australia for data linkage and the Australian Electoral Commission for establishing the initial sampling frame. We also acknowledge the contributions of the research assistants who helped to collate the data and manage the process, and we would like to thank the many thousands of Queensland residents who have participated in the QSkin Study (QSkin Investigators: David C. Whiteman, MBBS, PhD; Adele C. Green, MBBS, PhD; Catherine M. Olsen, PhD; Rachel E. Neale, B Vet Sc, PhD; Penelope M. Webb, D Phil. QSkin Study Team: Lea M. Jackman, Barbara A. Ranieri, Bridie S. Thompson, Rebekah A. Cicero). This work was supported by program grants from the National Health and Medical Research Council (NHMRC) of Australia (grant numbers 1073898 and 552429). DCW and REN are supported by NHMRC Research Fellowships. The funding body had no role in the design and conduct of the study, the collection, management, analysis, and interpretation of the data, or the preparation, review, or approval of the manuscript.

AUTHOR CONTRIBUTIONS

DCW was responsible for conception and design, interpretation of data, and drafting and revising the manuscript critically for important intellectual and statistical content. BST and APT conducted the statistical analyses and contributed to interpretation of data and drafting and reviewing the manuscript. M-CH and CM contributed to the statistical analyses. ACG, REN, and CMO contributed to the conception and design, interpretation of data, and critical revision of the manuscript for important intellectual content. All authors read and approved the final version of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at <http://dx.doi.org/10.1016/j.jid.2016.02.008>.

REFERENCES

- Cancer Council Australia, Position statement: screening and early detection of skin cancer. Cancer Council Australia; 2007.
- Chen AC, Martin AJ, Choy B, Fernandez-Penas P, Dalziel RA, McKenzie CA, et al. A phase 3 randomized trial of nicotinamide for skin-cancer chemoprevention. *New Engl J Med* 2015;373:1618–26.
- Choudhury K, Volkmer B, Greinert R, Christophers E, Breitbart EW. Effectiveness of skin cancer screening programmes. *Br J Dermatol* 2012;167(Suppl. 2):94–8.
- Eastham JA, May R, Robertson JL, Sartor O, Kattan MW. Development of a nomogram that predicts the probability of a positive prostate biopsy in men with an abnormal digital rectal examination and a prostate-specific antigen between 0 and 4 ng/mL. *Urology* 1999;54:709–13.
- English DR, Armstrong BK, Kricke A. Reproducibility of reported measurements of sun exposure in a case-control study. *Cancer Epidemiol Biomarkers Prev* 1998;7:857–63.
- Housman TS, Feldman SR, Williford PM, Fleischer AB, Goldman ND, Acostamadiedo JM, et al. Skin cancer is among the most costly of all cancers to treat for the Medicare population. *J Am Acad Dermatol* 2003;48:425–9.
- Katalinic A, Eisemann N, Waldmann A. Skin Cancer Screening in Germany. *Dtsch Arztebl Int* 2015;112:629–34.
- Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. *Br J Dermatol* 2012;166:1069–80.
- Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *Br Med J* 2009;338:b606.
- 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.
- Olsen CM, Green AC, Neale RE, Webb PM, Cicero RA, Jackman LM, et al. Cohort profile: the QSkin Sun and Health Study. *Int J Epidemiol* 2012;41:929–929i.
- Olsen CM, Neale RE, Green AC, Webb PM, Study QS, Epigene S, et al. Independent validation of six melanoma risk prediction models. *J Invest Dermatol* 2015;135:1377–84.
- Olsen CM, Williams PF, Whiteman DC. Turning the tide? Changes in treatment rates for keratinocyte cancers in Australia 2000 through 2011. *J Am Acad Dermatol* 2014;71:21–26.e1.
- Paul P, Pennell ML, Lemeshow S. Standardizing the power of the Hosmer–Lemeshow goodness of fit test in large data sets. *Stat Med* 2013;32:67–80.
- Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst* 2001;93:358–66.
- Stang A, Jockel KH. Does skin cancer screening save lives? A detailed analysis of mortality time trends in Schleswig-Holstein and Germany. *Cancer* 2015;122:432–7.
- Steffen A, MacInnis RJ, Joshy G, Giles GG, Banks E, Roder D. Development and validation of a risk score predicting risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2014;23:2543–52.
- Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Br Med J* 2009;338:b2393.
- Thompson BS, Olsen CM, Subramaniam P, Neale R, Whiteman DC. Medicare claims data reliably identify treatments for basal cell carcinoma and squamous cell carcinoma: a prospective cohort study [e-pub ahead of print]. *Aust New Zealand J Public Health* 2015; <http://dx.doi.org/10.1111/1753-6405.12478>.
- U.S. Department of Health and Human Services. The surgeon general's call to action to prevent skin cancer. Washington, DC: U.S. Department of Health and Human Services, Office of the Surgeon General; 2014.
- U.S. Preventive Services Task Force. Screening for skin cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Int Med* 2009;150:188–93.
- van der Mei IA, Blizzard L, Ponsonby AL, Dwyer T. Validity and reliability of adult recall of past sun exposure in a case-control study of multiple sclerosis. *Cancer Epidemiol Biomarkers Prev* 2006;15:1538–44.
- Vuong K, McGeechan K, Armstrong BK, Cust AE. Risk prediction models for incident primary cutaneous melanoma: a systematic review. *JAMA Dermatol* 2014;150:434–44.
- Whiteman D. Predicting melanoma risk: theory, practice, and future challenges. *Melanoma Management* 2014;1:105–14.
- Yu CL, Li Y, Freedman DM, Fears TR, Kwok R, Chodick G, et al. Assessment of lifetime cumulative sun exposure using a self-administered questionnaire: reliability of two approaches. *Cancer Epidemiol Biomarkers Prev* 2009;18:464–71.