"Repeatability of Repeatability": the stability of self-reported melanoma risk factors in two independent samples

Alexandra Mortimore,[1] Nirmala Pandeya,[1] Catherine M. Olsen,[1,2] David C. Whiteman[1,2]

Abstract

Objective: To determine the test-retest repeatability of a self-completed survey with items capturing skin cancer risk factors.

Methods: We invited 238 randomly selected participants of the QSkin II cohort to complete the baseline survey a second time. Responses were compared using kappa statistics and intraclass correlation coefficients to quantify agreement for categorical and continuous variables, respectively. We compared the performance of key items with that observed in an earlier repeatability study using the same survey instrument in an independent cohort.

Results: Measures of phenotypic characteristics had moderate to almost-perfect test-retest repeatability (e.g. eye colour weighted kappa (κ_w) = 0.87, 95% confidence interval [CI]: 0.81, 0.92). Items measuring sun exposure showed lower agreement (κ_w range 0.36-0.54) compared with phenotypic characteristics (κ_w range 0.59-0.87). Items relating to treatment of skin cancers demonstrated almost-perfect test-retest repeatability (e.g. excisions for skin cancers κ_w 0.85, 95%CI: 0.80, 0.89). In aggregate, the repeatability of key items was very similar across the two independent repeatability samples.

Conclusion: Fair to almost-perfect repeatability for self-reported skin cancer risk factors was robust across independent and temporally distant cohorts.

Implications for public health: These self-assessed risk factors for skin cancer are repeatable and suitable for use in clinical practice and research.

Key words: melanoma, skin cancer, survey, repeatability, validity

Skin cancer is Australia's 'national cancer'; this nation suffers the highest incidence of cutaneous melanoma and keratinocyte cancer in the world.[1-3] Ultraviolet radiation exposure is the principal environmental risk factor; established constitutional risk factors include fair skin, hair and eye colour, high numbers of melanocytic nevi and freckling tendency.[4,5]

Studies of skin cancer epidemiology, and efforts to monitor and evaluate skin cancer control activities require valid and reliable instruments for measuring these phenotypic and environmental risk factors. The evidence for reproducibility of self-reported measures of skin cancer risk factors is limited. In relation to constitutional melanoma risk factors, previous studies have demonstrated substantial agreement for eye colour, moderate to substantial agreement for skin and hair colour and fair to substantial agreement for skin phototype (burning / tanning), freckling and questions related to number of nevi.[6-16] Regarding environmental melanoma risk factors, previous studies reported fair to substantial agreement for sunburn history and sunscreen use, moderate agreement for hat use and time in the sun, and moderate to almost perfect agreement for sunbed use.[7-17] However, to date there have been only a handful of studies that have assessed the repeatability of important melanoma risk factors including nevus counts, sunscreen use on different parts of the body and family history of melanoma.[6-12,14-17] In addition, the stability of these estimates of repeatability remains unclear.

We have previously reported the repeatability of self-reported survey items for a population-based cohort (QSkin, recruited in 2011) that was established specifically for skin cancer research.[13] In that earlier investigation, we assessed the repeatability of self-reported items measuring phenotypic characteristics, past sun exposure, photoprotective behaviors, skin cancer history and treatment, and medical and social history. In 2019, we recruited a new cohort (QSkin II); this new cohort was a convenience sample of volunteers who completed a baseline survey and provided a saliva sample.

We sought to assess the test-retest repeatability of the same self-reported survey items that we used previously in a newly recruited, completely independent sample of participants. In so doing, we sought to compare the metrics of key items across the two different cohorts and thereby identify those self-reported risk factors demonstrating the most stable performance.

1. Cancer Control Group, QIMR Berghofer Medical Research Institute, Queensland
2. Faculty of Medicine, The University of Queensland, Queensland
Correspondence to: Professor David Whiteman, Cancer Control Group, QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane and Women's Hospital, QLD 4029; e-mail: david.whiteman@qimrberghofer.edu.au
Submitted: March 2021; Revision requested: June 2021; Accepted: June 2021
The authors have stated they have no conflicts of interest.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
Aust NZ J Public Health. 2021; 45:469-73; doi: 10.1111/1753-6405.13147

© 2021 The Authors

Methods

The QSkin study

The QSkin Sun and Health study is a cohort of Australian men and women established with the aim of prospectively investigating the role of genetic and environmental factors in developing cutaneous melanoma and other forms of skin cancer. The original cohort was recruited in 2010/2011 and comprised 43,794 residents of Queensland randomly selected from the Australian Electoral Roll and who completed a baseline survey. The cohort has been followed both passively and actively for melanoma and keratinocyte cancer outcomes.[18] In 2019 a new cohort was established from a convenience sample of Australian residents aged 18 years or older (n=8,656); participation involved completing the baseline survey and providing a saliva sample.

Study design

Eligible participants for this new repeatability study were those who had taken part in the QSkin II Study and returned their primary survey response between 29/01/2020 – 25/04/2020 and had completed 100% of the survey. Initially 500 participants were randomly selected from the QSkin II cohort with aim of receiving a completed second set of survey responses from at least 220 participants (i.e. an estimated re-completion fraction of 44%).

We sent a personalised email to each participant inviting them to complete the same QSkin survey for a second time. Each participant provided written informed consent, which was provided online via a link contained in the email. Involvement in this sub-study was voluntary and participants were able to withdraw at any time. Once consent was obtained, the online QSkin survey was accessed via a link embedded in the email invitation. Recruitment emails were sent in five waves over 29/1/20 to 6/2/20 in batches of 86 to 94. No reminder emails were sent to participants. In total, 440 invitations were sent until the target recruitment number of at least 220 completed surveys was reached. In total 238 agreed to take part (response fraction 54%).

The QSkin II Study and this repeatability sub-study were approved by the Human Research Ethics Committee at the QIMR Berghofer Medical Research Institute (approval number P3434).

Questionnaire

The QSkin survey was developed by the study investigators based on numerous similar instruments used in earlier studies of skin cancer.[6,19] The survey was designed to be self-administered and is publicly available here: https://www.qimrberghofer.edu.au/wp-content/uploads/2021/02/P3434_QSKIN_GENETICS_NewSample_Survey_V1.1_17Nov2018.pdf. Selected questions analysed in this repeatability study focused on risk factors for melanoma and other skin cancers including phenotype and sun exposure, in addition to medical history, family history and social history. The survey questions that assessed phenotype included skin color, skin burning and tanning type; eye and hair color; and various measures of freckling and nevus density. (Supplementary Table 1). Sun exposure assessments included number of severe sunburns in childhood, adolescence and adulthood; and number of hours spent in the sun on weekdays and weekends across different ages and frequency of sunbed or tanning bed use. We assessed sun protection behaviors through questions about use of sunscreen and hats. We also asked about medical history, including self-rated health, detailed questions about treatments for actinic lesions, and other aspects of medical and family history relating to skin cancers and melanoma. Finally, we captured information on height and weight, smoking, alcohol, other medical conditions and other aspects of socio-demographics.

Statistical analysis

We assessed the agreement between each participant's responses to the original survey and the repeat survey. We evaluated test-retest repeatability using simple Cohen's kappa (κ) for two level categorical variables, weighted Cohen's kappa (κ_w) for higher order categorical variables, and intraclass correlation coefficient (ICC) for continuous variables. Kappa values range from -1.0 to 1.0 with ≤0 indicating no agreement, 0.01–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial and 0.81–1.00 as almost perfect agreement.[20] The weighted kappa statistic compares the observed agreement with higher weights to responses with higher disagreement (or off diagonal responses). Intraclass coefficient (ICC) compares the repeatability of continuous measures taken repeatedly on the same subject. ICC values ≤0.5 indicate poor repeatability, values between 0.5–0.75 indicate moderate repeatability, 0.75–0.9 indicate good repeatability and[3] ≥0.90 indicate excellent repeatability.[21]

To assess the stability of item repeatability over time and across independent sampling frames, we compared the measures of agreement obtained in our first repeatability study[13] with those obtained in this new cohort.

Results

The average time interval between the first and second self-administered surveys in the QSkin II repeatability study was five months. Participants in the repeatability sample were similar to the QSkin II cohort for most characteristics, however, they were more likely to be male (52.1% vs. 35.3%; $p<0.001$), to have a history of skin cancers excised ($p=0.01$) and skin lesions frozen/burnt off ($p<0.001$) (Supplementary Table 2).

Of the self-reported risk factors for melanoma, we found items recording phenotypic characteristics including skin colour, skin phototype (burning and tanning), eye colour, hair colour, freckling at 21 years old and nevus count at the age of 21 years old to achieve moderate to almost perfect agreement in the QSkin II sample (κ_w 0.59–0.86) (Table 1). Self-reported number of nevi on the left upper arm (ICC 0.60) and total number of moles at age 21 (κ_w 0.59) demonstrated moderate repeatability. In contrast, the survey item measuring large nevi (>5mm) counts on the left upper arm showed poor repeatability (ICC 0.48; 95% CI 0.36–0.60).

We found items measuring sun exposure to have lower repeatability than phenotypic characteristics. The number of reported severe sunburns at various ages demonstrated moderate agreement (κ_w 0.44–0.54) while time spent in the sun at different ages had fair to moderate agreement (κ_w 0.36–0.51) (Table 1). Photoprotective behaviours were reasonably well reported, with weighted kappas for sunscreen use at different body sites ranging from 0.57 to 0.69; sunscreen application to the face had the highest repeatability of the sunscreen items (κ_w 0.69). Hat use when outdoors had substantial repeatability (κ_w 0.66), and history of sunbed use demonstrated almost perfect agreement between first and second surveys (κ_w 0.91) (Table 1).

© 2021 The Authors

Table 1. Repeatability of QSkin II survey items related to phenotype, sun exposure, and sun protection, medical and family history.

Survey items	N	Measure	Value	95% CI
Phenotype				
Skin colour	235	Weighted kappa	0.75	0.66, 0.84
Skin burning type	235	Weighted kappa	0.59	0.52, 0.67
Skin tanning type	234	Weighted kappa	0.59	0.51, 0.67
Eye colour	234	Weighted kappa	0.87	0.81, 0.92
Hair colour	234	Weighted kappa	0.86	0.80, 0.91
Grey hair now	232	Weighted kappa	0.82	0.77, 0.88
Grey hair start age	163	ICC	0.85	0.80, 0.88
Freckling on face, at 21yr	235	Weighted kappa	0.73	0.66, 0.79
Moles at age 21yr	234	Weighted kappa	0.59	0.51, 0.67
Moles left upper arm	149	ICC	0.60	0.51, 0.69
Large moles left upper arm (>5mm)	149	ICC	0.48	0.36, 0.60
Sun exposure				
Number of severe sunburns aged <10yr	222	Weighted kappa	0.54	0.47, 0.62
Number of severe sunburns aged 10-20yr	233	Weighted kappa	0.53	0.46, 0.61
Number of severe sunburns aged >20yr	231	Weighted kappa	0.44	0.34, 0.53
Time in the sun-weekdays (past year)	231	Weighted kappa	0.44	0.35, 0.53
Time in the sun-weekdays (aged 10-19yr)	331	Weighted kappa	0.36	0.27, 0.45
Time in the sun-weekdays (aged 20-29yr)	232	Weighted kappa	0.48	0.40, 0.56
Time in the sun-weekdays (aged 30-39yr)	229	Weighted kappa	0.44	0.36, 0.53
Time in the sun-weekend days (past year)	232	Weighted kappa	0.51	0.44, 0.59
Time in the sun-weekend days (10-19yr)	232	Weighted kappa	0.49	0.41, 0.57
Time in the sun-weekend days (20-29yr)	232	Weighted kappa	0.44	0.35, 0.53
Time in the sun-weekend days (30-39yr)	228	Weighted kappa	0.41	0.32, 0.49
Sunbed use	235	Weighted kappa	0.91	0.85, 0.98

Table 1 cont. Repeatability of QSkin II survey items related to phenotype, sun exposure, and sun protection, medical and family history.

Survey items	N	Measure	Value	95% CI
Sun exposure cont.				
Age at first use of sunbed	22	ICC	0.31	0.17, 0.51
Number of years lived in Northern Australia	110	ICC	0.78	0.71, 0.84
Number of years lived in Central Australia	159	ICC	0.89	0.86, 0.92
Number of years lived in Southern Australia	187	ICC	0.96	0.95, 0.97
Sun protection				
Routine (most days) sunscreen use - face	229	Kappa	0.69	0.59, 0.78
Routine sunscreen use – hands/ forearms	160	Kappa	0.63	0.48, 0.77
Routine sunscreen use – other body parts	129	Kappa	0.57	0.32, 0.82
Sunscreen use in the past year	235	Weighted kappa	0.59	0.51, 0.67
Hat use in the past year	233	Weighted kappa	0.66	0.58, 0.73
Medical history				
Self-rated health	234	Weighted kappa	0.64	0.57, 0.71
Skin cancers removed surgically	232	Weighted kappa	0.85	0.80, 0.89
Skin lesions burnt/frozen	232	Weighted kappa	0.86	0.82, 0.89
Perceived likelihood of getting melanoma	230	Weighted kappa	0.46	0.35, 0.56
Skin check by doctor (past 3yrs)	228	Weighted kappa	0.68	0.61, 0.75
Skin check by someone else (past 3yrs)	144	Weighted kappa	0.58	0.49, 0.68
Skin check by yourself (past 3yrs)	168	Weighted kappa	0.52	0.42, 0.62
Taking vitamin D supplementation	232	Kappa	0.76	0.67, 0.86
Diagnosed with diabetes	232	Weighted kappa	0.92	0.83, 1.00
Family history				
Family history of melanoma	231	Weighted kappa	0.70	0.61, 0.78
Strong family history of other cancer	232	Weighted kappa	0.45	0.34, 0.55
Are you a twin/twin in biological family	232	Kappa	0.76	0.66, 0.86

Notes:

Abbreviations: ICC = intraclass correlation coefficient; CI, confidence interval

Self-reported history of skin cancers treated with surgical excision demonstrated almost perfect agreement (κ_w 0.85), as did self-reported history of skin lesions burnt or frozen (κ_w 0.86). Items related to skin checks showed moderate agreement (κ_w 0.52-0.68) (Table 1). Family history of melanoma showed substantial repeatability (κ_w 0.70).

Questions relating to height, weight and work status had moderate to substantial test-retest repeatability with ICC 0.75, 0.99 and κ_w 0.87, respectively. Smoking status had almost perfect agreement (κ 0.89), and other measures of smoking history were also highly reproducible (Supplementary Table 3). We found that self-reports of alcohol intake had high repeatability (alcohol drinking status κ 0.81; alcohol weekly drinks κ_w 0.75). The survey item relating to use of cannabis demonstrated almost perfect agreement (κ_w 0.83), and medical conditions had moderate to high repeatability, except for hypertensive heart disease and eczema (Supplementary Table 4).

Comparison of item repeatability across cohorts

We compared the repeatability of survey items observed in the QSkin II cohort with those observed for the same items in the original QSkin cohort (Figures 1-3). The plots demonstrate remarkably consistent estimates of repeatability between the samples for all items. For phenotype items (Figure 1), measures of eye, hair and skin colour performed with very high repeatability in both samples. Measures of skin tanning and burning, and mole counts, were consistently less repeatable in both samples. Similarly, we found that items measuring sun exposure and sun protection (Figure 2) showed moderate agreement in both samples, except for sunbed use, which showed almost perfect agreement in both samples. The repeatability of items recording medical history and lifestyle factors (Figure 3) were comparable in the two samples, with recall of smoking status and skin cancer treatments consistently excellent, whereas recall of skin checks was consistently fair to moderate in both samples.

Discussion

We aimed to measure the test-retest repeatability of a self-completed questionnaire used for prospective research on melanoma and skin cancer and with items capturing skin cancer risk factors, sun exposure, photoprotective behaviors, medical and social history, and then to compare the

© 2021 The Authors

Figure 1: Comparison between the repeatability of phenotypic items as measured in the original QSkin cohort repeatability sample with the same items measured in the QSkin II repeatability sample.

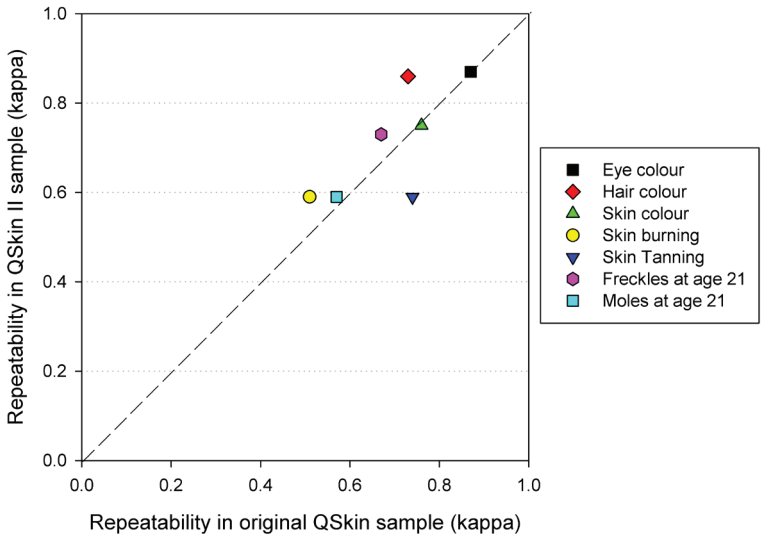


Figure 2: Correlation between the repeatability of measures of sun exposure, sun protection and sunbed use comparing the original QSkin cohort repeatability sample to the QSkin II repeatability sample.

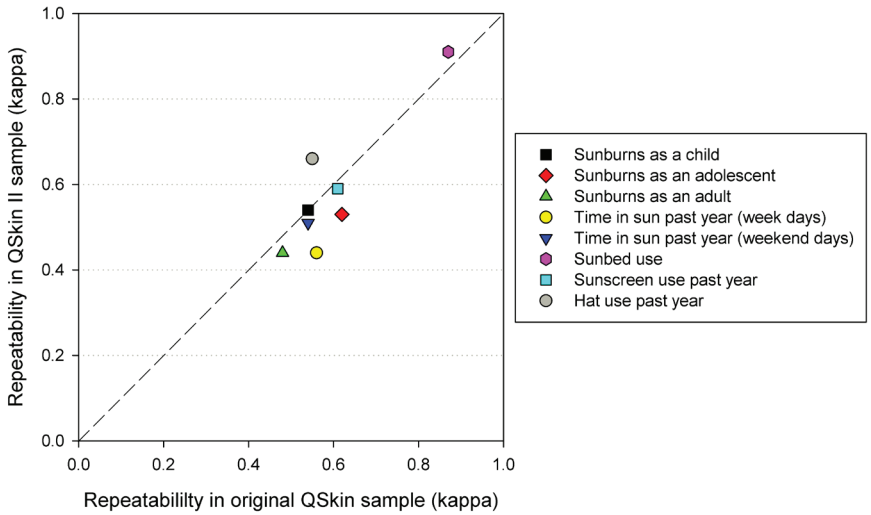
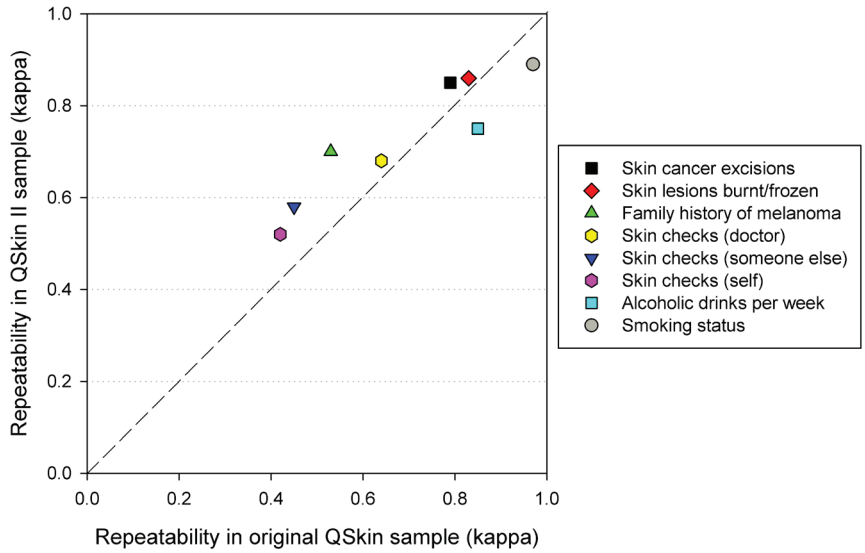


Figure 3: Correlation between the repeatability of items related to medical history and lifestyle factors comparing the original QSkin cohort repeatability sample to the QSkin II repeatability sample.



performance of key items across two cohorts recruited using different methods.

In general, we found very similar measures of repeatability in the second survey to those observed in the first. In other words, those items that performed well in the original QSkin cohort also performed well in the subsequent cohort. In general, measures of phenotype including eye colour and hair colour performed exceptionally well, while items relating to skin colour, skin phototype and freckling at 21 years old were also recalled well. We assessed three separate items relating to nevus burden: whole body nevus category at age 21 years, count of nevi >2mm on the left upper arm, and count of nevi >5mm on the left upper arm. Of these, the first two performed moderately well, while the third measure had limited repeatability. Whole body nevus category at 21 years is an influential factor that we and others have used in melanoma risk prediction models,[22] and our finding of good but not excellent repeatability highlights the challenge of developing prediction models based on imperfect measures. Although we found whole body nevus category to perform reasonably well, we may be at the limits of what self-report can achieve and therefore need to acknowledge that a degree of misclassification has to be accepted. Overall, aside from mole count on the left upper arm, the phenotypic variables had moderate to almost perfect repeatability, which underscores the confidence in using these items in melanoma risk prediction tools.

We found that measures of historical sun exposure (including severe sunburns and time in the sun) were less reliable than measure of phenotype, and thus more prone to misclassification. They remain challenging measures to improve as there are no 'gold standard' objective measures to validate them, and they rely completely on recall of the distant past. One alternative to quantifying sun exposure by recall is to use proxy measures such as history of sunspots. Our analyses show that these measures are recalled with very high repeatability. In the future, it may be possible to replace these recall-based methods with biomarkers that quantify cumulative DNA damage or other measures of actinic damage.[23] Interestingly, we observed that measures of sun protection performed well with moderate to substantial agreement, likely due to increased awareness of the hazards of sun exposure in Australia.

© 2021 The Authors

History of skin cancer treatment (whether by surgery or destructive means) performed exceptionally well, as did smoking status. Likewise, first degree family history of melanoma performed well, a reassuring finding given its frequent use in many skin cancer risk prediction tools.[22] History of skin checks by a doctor was also well recalled, especially when compared to the measures assessing skin checks conducted by self or someone else.

As for the QSkin surveys, previous studies assessing skin cancer risk factor repeatability have used an even mix of self-administered and face-to-face instruments.[6-12,14-17] Four of those earlier studies were conducted in Northern Hemisphere populations[10,11,14,15] and two studies took place in Australia.[6,13] The mean interval between the baseline survey and the repeatability survey in previous surveys ranged from three weeks up to 17 years;[6-12,14-17] in QSkin II the mean interval between surveys was five months. In general, our findings accord with those previously published measures of survey repeatability.

While most of our findings accord with others, we found higher levels of agreement for hat use when outdoors compared with the two previous studies reporting on this item.[10,13] It is possible that the generally higher prevalence of hat wearing in Queensland than other settings contributed to this higher repeatability.

A major strength of our study was the ability to compare the repeatability of survey items across two cohorts recruited in different ways and separated in time by more than six years.[13] The 'repeatability of the repeatability' was remarkably high for almost all items on the survey, suggesting that some items (e.g. history of treatments for actinic skin lesions; smoking; height; weight; some measures of skin phenotype) are always reported with very high consistency. Epidemiologic analyses based on these variables can therefore be accorded a high degree of confidence. Other items, such as measures of sun exposure and sun protection, are less reliable, and so epidemiologic inferences must be more cautious.

Participants were randomly selected from the QSkin II cohort to participate in the repeatability sub-study, reducing the chance of selection bias. Further, the selected participants were unaware at baseline that they would be asked to complete the survey a second time, minimising the chance of recall bias. The interval of five months between the baseline and repeated surveys is sufficiently long that respondents would be unlikely to precisely recall their previous responses to survey items, but is not so long as to allow substantial changes in exposure that would detract from an assessment of repeatability. While the repeatability sample was similar to the QSkin II cohort in terms of phenotype, photoprotective behaviours, smoking status and highest educational qualification, the repeatability sample had a higher proportion of males and more people reporting past history of treatments for actinic skin lesions. It is not obvious whether over-representation of these traits would systematically influence the consistency of self-reported assessments; we have no evidence to suspect so but cannot rule out the possibility.

In summary, questionnaires capturing information about risk factors for skin cancer demonstrate consistent levels of repeatability when administered to different community samples, suggesting that these self-reported instruments are reliable tools for measuring phenotypic and environmental risk factors for skin cancer.

Funding

This work was supported by National Health and Medical Research Council of Australia (grants numbers APP1073898, APP1058522). DCW is supported by a Research Fellowship from the NHMRC (APP1155413).

References

1. Whiteman DC, Green AC, Olsen CM. The growing burden of invasive melanoma: Projections of Incidence rates and numbers of new cases in six susceptible populations through 2031. *J Invest Dermatol*. 2016;136(6):1161-71.
2. Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. *Br J Dermatol*. 2012;166(5):1069-80.
3. Pandeya N, Olsen CM, Whiteman DC. The incidence and multiplicity rates of keratinocyte cancers in Australia. *Med J Aust*. 2017;207(8):339-43.
4. Schadendorf D, van Akkooi ACJ, Berking C, Griewank KG, Gutzmer R, Hauschild A, et al. Melanoma. *Lancet*. 2018;392(10151):971-84.
5. Savoye I, Olsen CM, Whiteman DC, Bijon A, Wald L, Dartois L, et al. Patterns of ultraviolet radiation exposure and skin cancer risk: The E3N-SunExp Study. *J Epidemiol*. 2018;28(1):27-33.
6. Baxter AJ, Hughes MC, Kvaskoff M, Siskind V, Shekar S, Aitken JF, et al. The Queensland STUDY OF MELANOMa: Environmental and Genetic Associations (Q-MEGA). Study design, baseline characteristics, and repeatability of phenotype and sun exposure measures. *Twin Res Hum Genet*. 2008;11(2):183-96.
7. Westerdahl J, Anderson H, Olsson H, Ingvar C. Reproducibility of a self-administered questionnaire for assessment of melanoma risk. *Int J Epidemiol*. 1996;25(2):245-51.
8. Berwick M, Chen YT. Reliability of reported sunburn history in a case-control study of cutaneous malignant melanoma. *Am J Epidemiol*. 1995;141(11):1033-7.
9. Bränström R, Kristjansson S, Ullén H, Brandberg Y. Stability of questionnaire items measuring behaviours, attitudes and stages of change related to sun exposure. *Melanoma Res*. 2002;12(5):513-19.
10. de Waal AC, van Rossum MM, Kiemeney LA, Aben KK. Reproducibility of self-reported melanoma risk factors in melanoma patients. *Melanoma Res*. 2014;24(6):592-601.
11. Glanz K, Schoenfeld E, Weinstock MA, Layi G, Kidd J, Shigaki DM. Development and reliability of a brief skin cancer risk assessment tool. *Cancer Detect Prev*. 2003;27(4):311-15.
12. McMullen EA, Dolan OM, McCarron P, Kee F. Reliability testing of a sun exposure questionnaire for the Northern Ireland population. *J Eur Acad Dermatol Venereol*. 2007;21(8):1071-3.
13. 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(12):1329-36.
14. Parr CL, Hjartåker A, Laake P, Lund E, Veierød MB. Recall bias in melanoma risk factors and measurement error effects: A nested case-control study within the Norwegian Women and Cancer Study. *Am J Epidemiol*. 2009;169(3):257-66.
15. Veierød MB, Parr CL, Lund E, Hjartåker A. Reproducibility of self-reported melanoma risk factors in a large cohort study of Norwegian women. *Melanoma Res*. 2008;18(1):1-9.
16. Rosso S, Minarro R, Schraub S, Tumino R, Franceschi S, Zanetti R. Reproducibility of skin characteristic measurements and reported sun exposure history. *Int J Epidemiol*. 2002;31(2):439-46.
17. English DR, Armstrong BK, Kricker A. Reproducibility of reported measurements of sun exposure in a case-control study. *Cancer Epidemiol Biomarkers Prev*. 1998;7(10):857-63.
18. 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(4):929-929i.
19. Green AC, Williams GM, Logan V, Strutton GM. Reduced melanoma after regular sunscreen use: Randomized trial follow-up. *J Clin Oncol*. 2011;29(3):257-63.
20. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-82.
21. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-63.
22. Vuong K, Armstrong BK, Weiderpass E, Lund E, Adami HO, Veierod MB, et al. Development and external validation of a melanoma risk prediction model based on self-assessed risk factors. *JAMA Dermatol*. 2016;152(8):889-96.
23. Olsen CM, Wilson LF, Green AC, Biswas N, Loyalka J, Whiteman DC. Prevention of DNA damage in human skin by topical sunscreens. *Photodermatol Photoimmunol Photomed*. 2017;33(3):135-42.

Supporting Information

Additional supporting information may be found in the online version of this article:

Supplementary Table 1: Phenotype, sun exposure, sun protection and medical history survey item responses.

Supplementary Table 2: Characteristics of 238 participants in the repeatability study compared with the entire cohort (QSkin II; n=7884).

Supplementary Table 3: Repeatability of QSkin height, weight, handedness, work and lifestyle survey items.

Supplementary Table 4: Repeatability of QSkin II medical conditions survey items.

© 2021 The Authors