Validity and Reliability of Adult Recall of Past Sun Exposure in a Case-Control Study of Multiple Sclerosis

I.A.F. van der Mei,[1] L. Blizzard,[1] A-L. Ponsonby,[2] and T. Dwyer[2]

[1]Menzies Research Institute, University of Tasmania, Hobart, Tasmania, Australia and [2]Murdoch Childrens Research Institute, Melbourne, Victoria, Australia

Abstract

Background: **Measurement of past sun exposure through recall by adults has the potential for measurement error. We aimed to investigate aspects of validity and reliability of self-reported past sun exposure.**

Methods: **A population-based case-control study was conducted in Tasmania on 136 cases with multiple sclerosis and 272 age- and sex-matched community controls. Repeat interviews on 52 cases and 52 controls were done on average 11 weeks after the initial interview. Sun exposure was assessed by questionnaire and lifetime calendar. Other measurements included serum 25-hydroxyvitamin D, actinic damage, and skin phenotype.**

Results: **There was an association between recent sun exposure and serum vitamin D (time in the sun: $r = 0.22$, $P < 0.01$; activities outside: $r = 0.31$, $P < 0.01$ for controls) and between lifetime sun exposure and actinic damage [correlation between 0.34 ($P < 0.01$) and 0.17 ($P = 0.01$) for controls]. The test-retest weighted κ statistic of self-reported sun exposure ranged from 0.43 to 0.74. Recall of childhood/adolescent sun exposure by standardized questioning was no less reproducible than recall of recent adult sun exposure and no less reliable when made with the calendar method. Comparing the questionnaire and calendar method, the measures of childhood/adolescent sun exposure had a similar predictive validity for multiple sclerosis.**

Conclusions: **The results of this study provide further evidence that adults are able to recall past sun exposure with shown validity and reliability and present information about the possible reasons for the good reliability of recalled sun exposure measures.** (Cancer Epidemiol Biomarkers Prev 2006;15(8):1538–44)

Introduction

Excess or insufficient UV radiation (UVR) and serum vitamin D have been implicated in the etiology of several diseases, including melanoma and nonmelanoma skin cancer, lymphomas, multiple sclerosis, type 1 diabetes, and prostate cancer (1-4). For some diseases, the quantum or circumstances of sun exposure as a child may be important. We recently showed this in a case-control study of multiple sclerosis, finding that sun exposure reduced the estimated relative risk of multiple sclerosis with the subjects at least risk being those with higher exposure during childhood and early adolescence (1). For melanoma, there is some evidence that risk is increased particularly by childhood sun exposure that results in burning (5).

Measurement of childhood sun exposure by questionnaire requires recall of past events possibly several decades earlier, with potential for measurement error. Studies that have investigated and compared the quantum, type, and determinants of this error in different instruments and approaches provide important information relevant to controlling and reducing it. Unfortunately, there have been few such studies to date and no study has assessed the validity of past sun exposure against both short-term (such as vitamin D status) and long-term (such as actinic damage) biomarkers for criterion validity. Vitamin D status can be assessed by serum 25-hydroxyvitamin D [25(OH)D]. It provides a measure of recent sun exposure because it has a half-life time of 1 to 2 months (6), and nearly all (90-100%) of the vitamin D in the human body is sourced from exposure to sunlight (7). Actinic damage of the skin, assessed from silicon casts of the skin of the hand, has previously been used as a measure of lifetime sun exposure. It has been found to be associated with lifetime sun exposure (8), residence in a high UVR location (9), outdoor occupations and leisure activities (10), solar keratosis (10, 11), and basal and squamous cell cancer (10). To measure sun exposure, we used not only standardized questions in a face-to-face interview but also a personal residential and working history. This history, which we refer to as a lifetime calendar, used memorable events as guideposts to assist recall. This life events approach has shown to be an effective strategy to record exposure histories accurately (12-14). A similar strategy was used to measure sun exposure in several other studies (4, 15-17).

Three validity studies have been conducted in school children (18-20), showing that these children are able to recall recent and current habitual sun exposure. In adolescents, polysulfone badge recordings of UVR exposure on 4 weekend days in late spring correlated well with two questions on habitual sun exposure (correlation 0.35 for time in the sun, 0.29 for activities outside; ref. 18). In 8-year-old (19) and adolescent (20) children, recent sun exposure was associated with levels of serum 25(OH)D. In adults, current sun exposure recorded in a diary has been shown to correlate with personal UVR dosimeters (18, 21), and lifetime sun exposure has been associated with actinic damage (8). Two studies assessed the reproducibility of recall of childhood and adolescent sun exposure by adults (8, 22). Rosso et al. (22) and English et al. (8) conducted retests after 18 to 26 months and 5 years, respectively, and observed a good level of agreement though slightly lower than that observed for lifetime sun exposure.

To broaden the range of findings available on this topic, we investigated in the Tasmanian Multiple Sclerosis case-control study the following: (*a*) the criterion validity of self-reported past sun exposure by comparing recent adult sun exposure with serum 25(OH)D status and by comparing lifetime sun

Received 12/26/05; revised 5/10/06; accepted 6/1/06.

Grant support: National Health and Medical Research Council of Australia, the Australian Rotary Health Research Fund, and MS Australia. I.A.F. van der Mei was supported by the Cooperative Research Centre for Discovery of Genes for Common Human Diseases.

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.

Requests for reprints: I.A.F. van der Mei, Menzies Research Institute, University of Tasmania, Private Bag 23, Hobart, Tasmania 7001, Australia. Phone: 61-3-6226-7700; Fax: 61-3-6226-7704. E-mail: Ingrid.vanderMei@utas.edu.au

doi:10.1158/1055-9965.EPI-05-0969

exposure with actinic damage, (*b*) the test-retest reproducibility of self-reported sun exposure during childhood and adolescence and compared it with that of recent adult sun exposure, and (*c*) the agreement between measurements by questionnaire and lifetime calendar of self-reported sun exposure during childhood and adolescence and compared it with that of recent adult sun exposure.

Materials and Methods

Overview. Validity refers to the degree to which a measurement measures what it purports to measure (23). Several types of validity can be distinguished, including construct validity, content validity, and criterion validity. Criterion validity examines the extent to which the measurement correlates with an external criterion of the phenomenon under study (23). In our study, we examined criterion validity by comparing self-reported recent sun exposure with serum 25(OH)D and by comparing self-reported lifetime sun exposure (from a lifetime calendar) with actinic damage.

Reliability refers to the reproducibility of a measure, that is, how consistently a measure can be repeated in the same subjects (23). We examined the intramethod reliability (or test-retest reproducibility) of standardized questions on sun exposure by measuring a subsample of participants at two points in time. We also examined the intermethod reliability of self-reported sun exposure by comparing standardized questions with a lifetime calendar method. For both the intramethod reliability and intermethod reliability, we compared recall of childhood/adolescence sun exposure with recall of recent sun exposure. Lastly, because sun exposure is associated with multiple sclerosis in this sample (1), we compared how well sun exposure in both methods (standard questionnaire and lifetime calendar) could statistically predict multiple sclerosis, noting that causality for this association has not been established (24).

Participants. The total sample consisted of 136 cases and 272 controls of the Tasmanian Multiple Sclerosis case-control study (1). Participants were persons ages <60 years who were residents of Tasmania, Australia and who had at least one grandparent born in Tasmania. Eligible cases had cerebral magnetic resonance imaging abnormalities consistent with multiple sclerosis (25) and clinically definite multiple sclerosis (26). Controls were selected from the comprehensive roll of electors registered in vote in Tasmanian elections. For each case, two control subjects were randomly selected and matched to the index case on sex and birth year. For the 136 cases included in the study, 272 eligible controls participated (participation rate, 76%). The subjects were interviewed between March 1999 and June 2001 by two research assistants: one in the north of the state (telephone district 63 and 64) and one in the south of the state (telephone district 62). Repeat interviews (T2) of 52 cases and 52 controls were conducted on average 11.1 weeks (SD, 2.9 weeks) after the initial interview (T1). The participants in T2 were sourced from the first 55 cases (participation rate, 95%) and 52 controls (participation rate, 100%) in the main study who lived in the south of Tasmania. A different interviewer conducted the T2. The project received ethics approval from the Human Research Ethics Committee of the Royal Hobart Hospital. Written consent was obtained from all participants.

Measurements during T1

Questionnaire. The interviewer-administered questionnaire sought information about the amount of sun exposure during four age spans of life (6-10 years, 11-15 years, 16-20 years, and the last 3 years) separately for summer and winter. Firstly, subjects were asked about their ''time in the sun'' with the question 'during weekends and holidays, how much time would you normally have spent in the sun in the following age spans?' and the response categories were <1, 1 to 2, 2 to 3, 3 to 4, and ≥4 hours daily. Secondly, they were asked about their ''activities outside'' with the question 'how much did your activities (playing, day sports, spectator sports, gardening, walking, working activities, etc.) take you outside in the following age spans?' and the response categories were not that often, a moderate amount, quite a lot, and virtually all the time. Occupational sun exposure was assessed with a question that asked whether their jobs were overall mainly indoors, both indoors and outdoors, or mainly outdoors. The questionnaire also included questions on lifetime sunburns (never, once, 2-5 times, 6-10 times, and >10 times) and sun-sensitive skin phenotype (tendency to burn: never burn, burn after >2-hour sun exposure, burn after 1 to 2 hours, burn after 0.5-1 hour, and burn within 0.5 hour; skin reaction: burn then peel, burn then tan, and tan only; ability to tan: a dark tan, a medium tan, light tan, and practically no tan).

Lifetime Calendar. Before T1, subjects were sent a lifetime calendar and asked to complete it for each year of their life. This required them to provide annual information on place of residence, school attended or type of work, and number of days weekly spent in each. During the interview, subjects answered the ''time in the sun'' question for summer only for each year of their life from 6 years to their current age using memorable events as guideposts to assist recall. The information already filled out on the calendar was used to identify blocks of years where the ''time in the sun'' lifestyle was constant or not. In addition, for each year of occupation, they were required to indicate how much time to the closest 15 minutes they would spend outside on a usual working day during working hours. For the comparison with the questionnaire data also taken at T1, an average of the rank score was taken of the annual calendar data for the relevant age spans.

Objective Measures. Serum 25(OH)D levels at interview were measured with a commercially available RIA (DiaSorin, Inc., Stillwater, MN) for 136 cases and 262 controls. Silicon skin surface casts of the hand, measuring actinic damage, were obtained as an objective marker of cumulative lifetime sun exposure (1). The grader was blinded to subject identity. Skin phenotype was assessed with a spectrophotometer at the upper inner arm and buttock, body sites usually not exposed to sunlight. Cutaneous melanin density was estimated from the skin reflectance of light centered at 400 and 420 nm (27).

Measurements during T2. T2 included all questions mentioned in the questionnaire section and the measurement of actinic damage using silicon casts. The lifetime calendar, serum 25(OH)D, and the assessment of melanin density with the spectrophotometer were not repeated.

Data Analysis

Assessment of Criterion Validity. To compare the association between serum 25(OH)D and questionnaire and calendar measures of recent adult exposure (both taken at T1), Pearson correlation coefficients were calculated. The questionnaire measures were ''time in the sun'' in summer and winter during the last 3 years and ''activities outside'' in summer and winter during the last 3 years. Because serum 25(OH)D is a marker of recent sun exposure, we also included correlations comparing 25(OH)D with sun exposure in summer for subjects who had their samples taken in summer or autumn (December-May) and with sun exposure in winter for those who had their samples taken in winter or spring (June-November).

Next, using Pearson correlations, we compared actinic damage with cumulative lifetime sun exposure measures obtained from the calendar. The analysis was restricted to 323 subjects with high quality casts, excluding 39 subjects whose casts were difficult to grade and 46 whose casts were

Table 1. Characteristics of subjects in the total sample at T1 and the sample that participated in the reliability study (T2)

Subject characteristics	Total sample at T1		Repeat sample at T2	
	Cases (n = 136)	Controls (n = 272)	Cases (n = 52)	Controls (n = 52)
Female/Male ratio	2.09	2.09	1.74	2.25
Age (y), mean (SD)	44.0 (9.2)	44.4 (9.2)	43.6 (10.6)	44.7 (10.0)
Height (cm), mean (SD)	166.9 (9.0)	166.1 (8.9)	167.0 (8.5)	165.6 (7.7)
Weight (kg), mean (SD)	72.9 (14.4)	76.2 (16.1)	73.2 (14.4)	71.7 (12.5)
Born in Tasmania (%)	96.3	95.2	96.2	93.9
Living in Tasmania 10 years of age (%)	95.6	95.2	98.1	93.9
Age at diagnosis (y), mean (SD)	34.8 (9.1)	—	35.5 (10.3)	—
Age at first symptoms (y), mean (SD)	32.2 (9.1)	—	33.3 (10.3)	—
Time since diagnosis (y), mean (SD)	9.4 (7.5)	—	8.7 (6.3)	—
Time since first symptoms (y), mean (SD)	12.0 (8.0)	—	10.9 (6.3)	—
Expanded disability status scale score, mean (SD)	3.6 (2.2)	—	3.4 (2.4)	—
Type of multiple sclerosis				
Relapsing remitting (%)	65.4	—	64.7	—
Secondary progressive (%)	26.5	—	23.5	—
Primary progressive (%)	8.1	—	11.8	—

unable to be graded. The 85 excluded subjects did not differ from the 323 included subjects in terms of the characteristics reported in Table 1 or the indices of agreement and validity reported for other analyses. Including the data for the 39 subjects whose casts were difficult to grade did not influence the results of this analysis. Lifetime sun exposure during leisure days was calculated by aggregating ''time in the sun'' in summer during leisure days for each year of life between age 6 years and their current age after the categories of the question were assigned the following scores: <1 hour daily = 0.5 hour, 1 to 2 hours daily = 1.5 hours, 2 to 3 hours daily = 2.5 hours, 3 to 4 hours daily = 3.5 hours, and >4 hours daily = 4.5 hours. Lifetime sun exposure during work days was calculated by aggregating the time daily reportedly spent outside during work hours firstly over work days in the week, then over weeks in the year, and finally over years to date of working life. Total lifetime sun exposure was the aggregate of these.

Assessment of Test-Retest Reliability. The test-retest comparison between T1 and T2 was conducted on 52 cases and 52 controls. The proportion in exact agreement (percent agreement) and the weighted Cohen's κ statistic (28) were calculated as measures of agreement (29). To assess systematic underreporting or overreporting, Bowker's test of symmetry (30) was applied to the misclassification matrices (the cross-tabulations of the two measures). To investigate factors associated with agreement, we used log binomial regression to model the probability of agreement (agreement = 1/disagreement = 0) as a function of predictors, such as question type, life span, and season of the year, using the generalized estimating equations algorithm to take account of correlated responses for the same individual. To assess confounding and effect modification by case/control status, age at interview, sex, education level, type of multiple sclerosis, disease duration, or disability level (assessed by expanded disability status scale), we entered linear predictors and interaction terms for these variables.

Assessment of Intermethod Reliability. The intermethod comparison of measurements by questionnaire and lifetime calendar (both taken at T1) was conducted on the total sample of 136 cases and 272 controls. To compare the questionnaire data for four times (6-10 years, 11-15 years, 16-20 years, and the last 3 years) with the annual measurements by lifetime calendar, we averaged the category rank scores of the annual calendar data for relevant years and rounded the result. The methods of analysis of test-retest reproducibility were used here also.

Intermethod Comparison of the Predictive Validity of Measures of Sun Exposure. The association between summer sun exposure and multiple sclerosis was assessed using unadjusted age- and sex-matched odds ratios and 95% confidence intervals (95% CI) estimated using conditional logistic regression. Results are presented for the ''time in the sun'' question, with responses dichotomized at 2 to 3 hours daily, comparing questionnaire and calendar methods.

Results

The sample of subjects at T2 was similar in most characteristics to the total sample of subjects (Table 1). A very high percentage of participants were born and living in Tasmania when 10 years of age (Table 1).

Criterion Validity of Measures of Self-Reported Past Sun Exposure

Association Between Recent Adult Sun Exposure and Serum 25(OH)D. We compared serum 25(OH)D with recent sun exposure indices. The results are shown in Table 2. Importantly,

Table 2. Correlations between serum 25(OH)D levels and self-reported sun exposure at T1

	Cases (n = 136)		Controls (n = 262)	
	Correlation	P	Correlation	P
Time in the sun in summer in the last 3 years	0.24	<0.01	0.20	<0.01
Time in the sun in winter in the last 3 years	0.19	0.03	0.16	<0.01
Time in the sun in the last 3 years *	0.31	<0.01	0.22	<0.01
Activities outside in summer in the last 3 years	0.30	<0.01	0.22	<0.01
Activities outside in winter in the last 3 years	0.19	0.03	0.18	<0.01
Activities outside in the last 3 years*	0.39	<0.01	0.31	<0.01
Time in the sun in summer in the last 3 years (calendar)	0.23	<0.01	0.20	<0.01
Time in the sun in summer in the last year (calendar)	0.22	<0.01	0.21	<0.01

*Sun exposure in summer for subjects who had their serum sample taken in the summer and autumn months and sun exposure in winter for subjects who had their serum sample taken in the winter and spring months.

Table 3. Correlations between actinic damage and lifetime sun exposure measures obtained from the calendar at T1

	Cases (n = 105)				Controls (n = 218)			
	r	P	Adjusted*		r	P	Adjusted*	
			r	P			r	P
Age	0.52	<0.01			0.43	<0.01		
Melanin at the upper inner arm (%)	−0.05	0.64			−0.21	<0.01		
Lifetime sun exposure, leisure days	0.35	<0.01	0.13	0.19	0.31	<0.01	0.08	0.26
Lifetime sun exposure, work days	0.22	0.03	0.17	0.08	0.26	<0.01	0.16	0.02
Total lifetime sun exposure	0.33	<0.01	0.20	0.05	0.34	<0.01	0.17	0.01

*Adjusted for age.

the correlations were much higher if the measurements of 25(OH)D and sun exposure were aligned correctly by season. The associations between time in the sun in summer in the last 3 years and 25(OH)D were nearly identical for the questionnaire and calendar-based method.

Association Between Lifetime Sun Exposure and Actinic Damage. We next compared actinic damage with three measures of cumulative sun exposure obtained from the calendar: lifetime sun exposure during leisure days, lifetime sun exposure during workdays, and total lifetime sun exposure during both leisure and workdays (Table 3). Each was strongly associated with actinic damage. Age was also strongly associated with actinic damage partly because of its strong association with the three measures of cumulative sun exposure (for controls: lifetime sun exposure leisure days, r = 0.60; lifetime sun exposure work days, r = 0.28; total lifetime sun exposure, r = 0.49) and partly because aging might have an effect on actinic damage that is independent of sun exposure. Higher melanin density was associated with lower actinic damage among controls. However, adjustment for melanin density at the upper inner arm had little effect. After adjustment for sun exposure after age 21, the association between sun exposure before age 21 and actinic damage was not significant (P = 0.15), reflecting partly the high correlation between sun exposure before and after age 21 (r = 0.46, P < 0.01).

Test-Retest Reliability of Reported Sun Exposure during Childhood and Adolescence. On retest after 11 weeks, the highest agreement for cases and controls alike was for the question on occupational sun exposure that had only three response categories (Table 4). For the other questions, we modeled the probability of exact agreement between the first and the second reports using log binomial regression. The proportion of responses in exact agreement seemed to vary with respect to the span of life referred to (6-10 years, 11-15 years, 16-20 years, and the last 3 years), generally being higher to questions on more recent exposure for cases (P < 0.01) but lower to questions on more recent exposure for controls (P = 0.03). However, these associations did not persist after taking into account whether participants provided an answer in an extreme category (cases, P = 0.68; controls, P = 0.55). People who on the first occasion gave an answer in an extreme category were more likely to provide the same answer on the second occasion compared with people who did not give an answer in an extreme category (cases, P < 0.01; controls, P = 0.01). For example, for ''time in the sun in summer,'' the number of controls (N = 52) who provided an answer in the extreme category ''>4 hours daily'' at the first interview decreased with recency of exposure: n = 31 for age span 6 to 10 years, n = 28 for 11 to 15 years, n = 17 for 16 to 20 years, and n = 11 for the last 3 years. Similarly, the number of people in exact agreement on the first and second occasion decreased with recency of exposure: n = 31, n = 30, n = 24, and n = 22 for age spans 6 to 10 years, 11 to 15 years, 16 to 20 years, and the last 3 years, respectively. With account taken in this way of the response distribution, agreement did also not vary by type of question (''time in the sun'' and ''time outside''; cases, P = 0.22 and controls, P = 0.13) or season of reported exposure (summer and winter; cases, P = 0.91 and controls, P = 0.96). Thus, the differences in proportions in exact agreement between questions in Table 4 were explained by differences in the proportion of responses in extreme categories.

Examination of the misclassification matrices produced no evidence that subjects systematically underreported or overreported at the second interview. The probability of agreement between the first and the second reports did not differ systematically by age at interview (cases, P = 0.67; controls, P = 0.28), sex (cases, P = 0.61; controls, P = 0.07), or education level (cases, P = 0.34; controls P = 0.10). Among cases, disagreement was not influenced by type of multiple sclerosis (P = 0.21) and disease duration (P = 0.93). Surprisingly, agreement was better among cases with a high disability (expanded disability status scale score; P = 0.02). This effect was slightly stronger for recent sun exposure and seemed to be explained that by the fact that cases with a high disability provided more often an answer in the lowest exposure category.

Table 4. Test-retest reliability between T1 and T2 (after 11 weeks) for self-reports of sun exposure

Question, age span (y)	Cases (n = 52)		Controls (n = 52)	
	a/N*	κ_w (95% CI)†	a/N*	κ_w (95% CI)†
Time in the sun in summer (five levels: <1 hour daily to >4 hours daily)				
6-10	26/52	0.65(0.38-0.91)	31/52	0.61(0.34-0.88)
11-15	25/52	0.59(0.35-0.84)	30/52	0.57(0.30-0.84)
16-20	17/52	0.57(0.31-0.83)	24/52	0.47(0.20-0.74)
Last 3	31/52	0.71(0.44-0.98)	22/51	0.60(0.33-0.86)
Time in the sun in winter (five levels: <1 hour daily to >4 hours daily)				
6-10	26/52	0.60(0.33-0.87)	24/52	0.65(0.38-0.92)
11-15	20/52	0.50(0.23-0.77)	24/52	0.43(0.18-0.69)
16-20	23/52	0.68(0.41-0.95)	22/52	0.56(0.29-0.82)
Last 3	34/52	0.74(0.47-1.00)	24/52	0.61(0.34-0.88)
Activities outside in summer (four levels: not that often to virtually all the time)				
6-10	26/52	0.58(0.31-0.85)	33/52	0.51(0.24-0.78)
11-15	25/52	0.58(0.31-0.84)	33/52	0.63(0.36-0.90)
16-20	28/52	0.63(0.36-0.90)	30/52	0.56(0.29-0.82)
Last 3	30/52	0.59(0.32-0.86)	26/52	0.52(0.25-0.79)
Activities outside in winter (four levels: not that often to virtually all the time)				
6-10	24/52	0.53(0.26-0.80)	28/52	0.61(0.34-0.88)
11-15	24/52	0.59(0.32-0.87)	28/52	0.70(0.43-0.97)
16-20	31/52	0.69(0.41-0.96)	31/52	0.62(0.35-0.88)
Last 3	37/52	0.63(0.36-0.90)	21/51	0.56(0.29-0.83)
Occupational sun exposure (three levels: mainly indoors to mainly outdoors)				
Working years	44/50	0.86(0.58-1.00)	45/52	0.85(0.58-1.00)

*Number of subjects with exact agreement on both reports/total number of subjects.
†Weighted κ statistic (95% CI).

Table 5. Test-retest reliability between T1 and T2 (after 11 weeks) for self-reports of sun-sensitive skin phenotype and measurements of actinic damage

Measure	Cases (n = 52)		Controls (n = 52)	
	a/N*	κ_w (95% CI) †	a/N*	κ_w (95% CI) †
Tendency to burn (five levels: never burn to burn within 0.5 h)	23/52	0.62(0.35-0.88)	23/52	0.65(0.38-0.91)
Skin reaction (three levels: burn then peel to tan only)	30/52	0.52(0.25-0.79)	41/52	0.76(0.49-1.00)
Ability to tan (four levels: dark tan to practically no tan)	37/52	0.75(0.48-1.00)	38/52	0.76(0.49-1.00)
Lifetime burns (five levels: never to >10 times)	28/51	0.73(0.45-1.00)	28/52	0.61(0.34-0.87)
Actinic damage (four levels: grade 3-6)	18/35	0.72(0.39-1.00)	34/43	0.88(0.58-1.00)

*Number of subjects with exact agreement on both reports/total number of subjects.
†Weighted κ statistic (95% CI).

For comparison, we show in Table 5 the test-retest results for self-reports of sun-sensitive skin phenotype. In analyses of responses by individuals, the skin questions overall had better agreement on retest than questions on childhood/adolescent sun exposure (cases, $P = 0.03$; controls, $P = 0.04$). This was principally due to high agreement on the ''tanning'' question (cases, $P < 0.01$; controls, $P < 0.01$) and, for controls ($P < 0.01$), on the ''skin reaction'' question. Also shown in Table 5 are the test-retest results for the assessments of actinic damage made using standardized laboratory protocols from silicon casts of subjects' hands. The probability of agreement was higher than that of the two sun questions for controls ($P < 0.01$) and similar for cases ($P = 0.85$) and nearly as high as the agreement for the question on occupational sun exposure.

Intermethod Reliability of Reported Sun Exposure during Childhood and Adolescence. We compared the results for all subjects of ''time in the sun'' in summer measured using the questionnaire and the lifetime calendar (Table 6). Again, the probability of exact agreement between the two measurements decreased with increasing recency of exposure (cases, $P < 0.01$; controls, $P < 0.01$), but this time, it persisted after adjustment for whether the questionnaire response was in an extreme category (cases, $P < 0.01$; controls, $P < 0.01$). This reflected a systematic trend of reporting higher exposure during more recent life spans by the calendar method. The changing distribution of measurements for controls is summarized in Table 7. The same pattern was observed for both interviewers (data not shown) and for cases (data not shown).

The probability of agreement for the ''time in the sun'' in summer between the questionnaire and the lifetime calendar method was similar for cases and controls ($P = 0.35$). Agreement was higher for women than for men when exposure was not in a high category ($P < 0.01$) and when recall of exposure at young ages was required ($P < 0.01$). Agreement between the two measurements did not differ by age ($P = 0.84$), education level ($P = 0.91$), type of multiple sclerosis ($P = 0.43$), or disease duration ($P = 0.87$) for cases. Similarly to the test-retest agreement, here, agreement was also better among cases with a high disability (expanded disability status scale score; $P < 0.01$). Again, this effect was slightly stronger for recent sun exposure and seemed to be explained that by the fact that cases with a high disability provided more often an answer in the lowest exposure category.

Table 6. Agreement between questionnaire and calendar measurements of time in the sun in summer during weekends and holidays at T1

Age span (y)	Cases (n = 136)		Controls (n = 268)	
	a/N*	κ_w (95% CI) †	a/N*	κ_w (95% CI) †
Time in the sun in summer (five levels: <1 hour daily to >4 hours daily)				
6-10	90/136	0.71(0.55-0.88)	181/268	0.58(0.47-0.70)
11-15	81/136	0.65(0.48-0.82)	169/272	0.61(0.49-0.72)
16-20	55/136	0.57(0.40-0.74)	127/271	0.58(0.46-0.70)
Last 3	49/136	0.55(0.40-0.70)	104/270	0.54(0.43-0.65)

*Number of subjects with exact agreement on both reports/total umber of subjects.
†Weighted κ statistic (95% CI).

Intermethod Comparison of the Predictive Validity of Measures of Sun Exposure. Lastly, we compared the association between multiple sclerosis and sun exposure (''time in the sun'' in summer) measured by the questionnaire and the calendar. For the questionnaire, this association was estimated for the four age spans (6-10 years, 11-15 years, 16-20 years, and the last 3 years). For the calendar, this association was estimated at every age from 6 to 20 years, at the current age of subjects, and for each year in the 3 years before their current age (Fig. 1). The odds ratios for multiple sclerosis for measurements of sun exposure made using the calendar were similar to those for measurements for the corresponding age span made using the questionnaire in at least three of the four age spans (Fig. 1).

Discussion

In this study, we provide evidence of the criterion validity of past sun exposure by self-report because sun exposure indices, used to measure sun in the last 3 years and lifetime sun exposure, were associated with biological markers of recent and lifetime sun exposure. We also showed that recall of childhood/adolescent sun exposure by standardized questioning is no less reproducible than recall of recent adult sun exposure and that the measurements were no less reliable when made with the calendar-based method that is increasingly being used in studies that measure sun exposure (1, 21, 31).

The sample of subjects for the reproducibility study was small (52 cases and 52 controls), which could have reduced the ability of this study to detect differences in reliability between subgroups. However, the repeat sample was similar in structure to the total sample of subjects and, reducing the potential for nonresponse bias, participation rates for both T1 and T2 were high. The use of a different interviewer for the repeat interviewers might have increased the variability and underestimated the true reliability of self-reported sun exposure, but the use of a standardized questionnaire and interview protocol limited the influence of the interviewer. The use of past ''time in the sun'' measures could have lead to substantial misclassification of the measurement of past sun exposure if participants had resided in locations with varying ambient UVR levels, but a very high proportion of participants lived in Tasmania for most of their lives and their estimated sun exposure would not be confounded by their residential history. We did not conduct a test-retest comparison using the calendar approach, but English et al. (8) found good agreement for this method. This is the only study that was able to assess criterion validity using both actinic damage and serum vitamin D levels.

Table 7. Comparison of measurements by questionnaire and lifetime calendar for controls of hours spent in the sun in summer at T1

Age span (y)	Response by calendar relative to questionnaire					
	2+ lower % (*n*)	1 category lower % (*n*)	Same category % (*n*)	1 category higher % (*n*)	2+ higher % (*n*)	Bowker's test of symmetry (*P*)
6-10	3.0 (8)	6.0 (16)	67.5 (181)	14.9 (40)	8.6 (23)	< 0.01
11-15	2.9 (8)	10.3 (28)	62.1 (169)	17.3 (47)	7.4 (20)	0.02
16-20	7.4 (20)	18.1 (49)	46.9 (127)	19.2 (52)	8.5 (23)	0.56
Last 3	4.8 (13)	8.9 (24)	38.5 (104)	25.6 (69)	22.2 (60)	< 0.01

The instruments used to measure past sun exposure seem to provide valid estimates when used to measure recent adult exposure or lifetime exposure. We found that sun exposure in the last year or 3 years correlated relatively well with vitamin D status, given there was an attenuation of the correlation due to measurement error, because sun exposure in the last year or 3 years was measured rather than sun exposure in the last 1 to 2 months. Lifetime sun exposure measured by calendar also correlated relatively well with actinic damage. It predicted ~10% of the variation of actinic damage, a commonly used criterion for quantitative significance (32). Adjustment for age in Table 1 may have resulted in an overadjustment because it was highly correlated with, and integral to, lifetime actinic damage. Therefore, the true association lies somewhere between the unadjusted and adjusted correlation coefficient.

These findings are consistent with those in other studies using samples with varied age groups. Two Tasmanian studies of 8-year-old (19) and adolescent (20) children found that serum 25(OH)D was correlated with usual levels of sun exposure in the season preceding the measurements [$r = 0.20$, $P < 0.01$ (19) and $r = 0.23$, $P = 0.01$ (20) for exposure during school holidays; $r = 0.16$, $P = 0.02$ (19) and $r = 0.23$, $P = 0.01$ (20) for exposure during weekends]. Lifetime sun exposure in adults measured by a calendar method was associated with actinic damage in the study by English et al. (8). Sun exposure assessed using a questionnaire in adolescents (18) or a diary method in adults (21) has also been shown to be correlated with short-term UVR exposure measured by polysulfone badges or personal UVR dosimeters.

We found that recall of childhood/adolescent sun exposure by standardized questioning was no less reproducible than recall of recent adult sun exposure. Two other reliability studies, examining agreement of sun exposure, also found that sun exposure in the distant past was recalled with reliability comparable with that of exposure in the recent past (8, 22). On retest after 5 years using a calendar method, English et al. (8) found an intraclass correlation coefficient of 0.55 (0.43-0.65) for sun exposure at ages 8 to 14 years and 0.77 (0.70-0.83) at ages 15 to 19 years. Their reliability coefficients for more recent exposure were 0.74 (0.66-0.80) at ages 25 to 34 years and 0.73 (0.65-0.75) at ages 35 to 39 years. On retest after 18 to 26 months, Rosso et al. (22) found an intraclass correlation coefficient of 0.65 (0.56-0.73) for the number of weighted outdoor hours at the beach in childhood. Their reliability coefficient for lifetime weighted outdoor hours at the beach was 0.79 (0.69-0.88). Thus, there are now three test-retest studies with different times between the interviews, showing the same finding that recall of sun exposure in childhood and adolescence is reproducible.

In this study, a major factor contributing to the high reproducibility of questions on childhood/adolescent sun exposure was the distribution of responses. Agreement on retest for controls was highest on questions on childhood exposure, but this was because a higher proportion of their responses to those questions fell in the highest exposure category, and those who reported the highest level of exposure at T1 tended to report the same category of exposure on the second occasion. The age span of life for which exposure was being recalled did not contribute to test-retest agreement independently of this factor. We suppose that disagreement would have been greater had we provided a wider choice of high exposure categories. In that sense, the relatively high level of agreement we found for childhood exposure was due in part to the structure of the questions. The reproducibility of

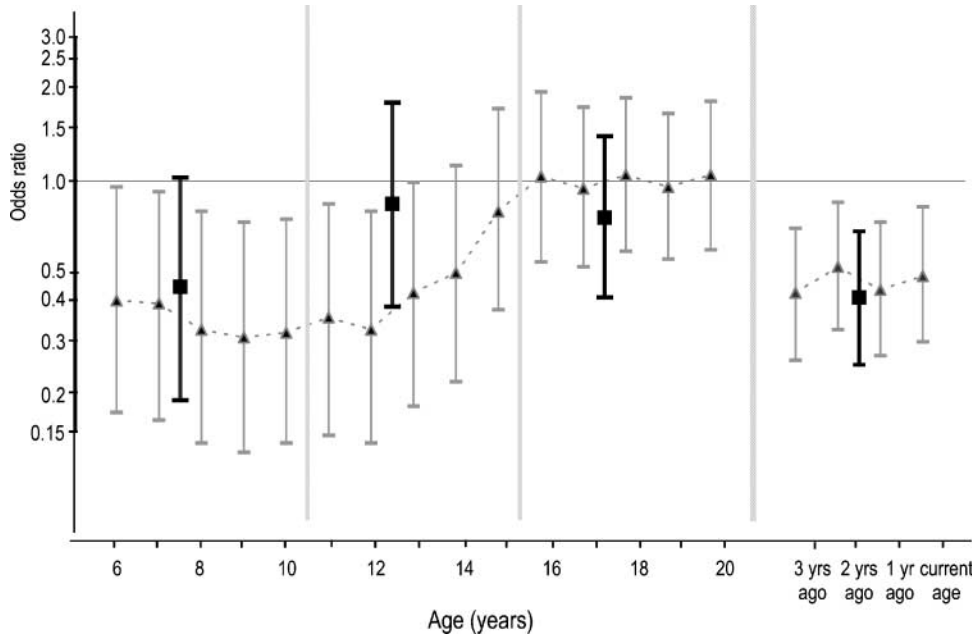

Figure 1. The association between summer sun exposure and multiple sclerosis by age. Unadjusted odds ratios and 95% CI for higher (≥2-3 hours daily on average) sun exposure in summer on weekends and holidays estimated by a lifetime calendar at each year of age between 6 and 20 years, at their current age, and at 1, 2, and 3 years before their current age and estimated by questionnaire for the age spans 6 to 10 years, 11 to 15 years, 16 to 20 years, and in the last 3 years.

recall of childhood/adolescent sun exposure was lower than that for some (but not all) questions on skin phenotype (33-36), a factor that does not change by stage of life, season, or weather. The reproducibility of recall of sun exposure was not influenced by factors, such as level of education, age, sex, case-control status, type of multiple sclerosis, disease duration, or disability level. A higher disability was associated with a higher agreement, but again, this seemed to be explained by the distribution of responses because cases with a high disability more often reported an answer in the lowest exposure category.

The life events calendar approach has been used in recent studies (1, 4, 15-17) to obtain information on sun exposure. The calendar approach provides flexibility to focus on exposure during specific periods of life and the information on residential and working history that is necessary to base an assessment of ambient UVR levels. A recent study showed that a shortened version of a calendar, administered via a telephone interview, had construct validity against the ''whole-of-life'' calendar [intraclass correlation coefficient, 0.65 (0.48-0.78); ref. 37]. Based on analysis of the ''whole-of-life'' calendar, this shortened version asks about sun exposure at the four most informative ages (10, 20, 30, and 40 years; ref. 37). The interview part of a lifetime calendar is less standardized, with more nonscripted interaction between interviewer and participant, and interviewers need to be trained carefully to limit interviewer administration effects. In this study, no evidence was found of interviewer effects but the measurements by calendar were systematically higher than the measurements by questionnaire, and this difference became increasingly pronounced for more recent exposure. This trend in systematic error coincided with diminished agreement between questionnaire- and calendar-based measurements even after adjustment for the distribution of responses. Although the intermethod agreement decreased for more recent sun exposure, there was no difference in the correlation between sun exposure in the last 3 years and vitamin D status between the questionnaire and calendar-based method. In addition, both methods showed a similar association between sun exposure and multiple sclerosis in three of the four age spans that were examined.

In conclusion, the results of this study provide further evidence that adults are able to recall past sun exposure with shown validity and reliability and present information about the possible reasons for the good reliability of recalled sun exposure measures.

Acknowledgments

We thank the participants and Trish Groom and Jane Pittaway for conducting the interviews, Natasha Newton for administrative support and data entry, Sue Sawbridge and Tim Albion for the development and management of the database, the Tasmanian Multiple Sclerosis Society for assisting with the recruitment of volunteers, and H Butkueven, A Hughes, B Drulovis, and S Sjieka who were involved with the clinical diagnosis.

References

1. van der Mei IA, Ponsonby AL, Dwyer T, et al. Past exposure to sun, skin phenotype, and risk of multiple sclerosis: case-control study. BMJ 2003;327:316.
2. Luscombe CJ, Fryer AA, French ME, et al. Exposure to ultraviolet radiation: association with susceptibility and age at presentation with prostate cancer. Lancet 2001;358:641–2.
3. Hypponen E, Laara E, Reunanen A, Jarvelin MR, Virtanen SM. Intake of vitamin D and risk of type 1 diabetes: a birth-cohort study. Lancet 2001;358:1500–3.
4. Hughes AM, Armstrong BK, Vajdic CM, et al. Sun exposure may protect against non-Hodgkin lymphoma: a case-control study. Int J Cancer 2004;112:865–71.
5. Elwood JM, Jopson J. Melanoma and sun exposure: an overview of published studies. Int J Cancer 1997;73:198–203.
6. Vieth R. Vitamin D supplementation, 25-hydroxyvitamin D concentrations, and safety. Am J Clin Nutr 1999;69:842–56.
7. Holick MF. Vitamin D: the underappreciated D-lightful hormone that is important for skeletal and cellular health. Curr Opin Endocrinol Diabetes 2002;9:87–98.
8. English DR, Armstrong BK, Kricker A. Reproducibility of reported measurements of sun exposure in a case-control study. Cancer Epidemiol Biomarkers Prev 1998;7:857–63.
9. Fritschi L, Green A. Sun damage in teenagers' skin. Aust J Public Health 1995;19:383–6.
10. Green AC. Premature ageing of the skin in a Queensland population. Med J Aust 1991;155:473–4, 477–8.
11. Holman CD, Evans PR, Lumsden GJ, Armstrong BK. The determinants of actinic skin damage: problems of confounding among environmental and constitutional variables. Am J Epidemiol 1984;120:414–22.
12. Hoppin JA, Tolbert PE, Flagg EW, Blair A, Zahm SH. Use of a life events calendar approach to elicit occupational history from farmers. Am J Ind Med 1998;34:470–6.
13. Friedenreich CM, Courneya KS, Bryant HE. The lifetime total physical activity questionnaire: development and reliability. Med Sci Sports Exerc 1998;30:266–74.
14. Krall EA, Valadian I, Dwyer JT, Gardner J. Accuracy of recalled smoking data. Am J Public Health 1989;79:200–2.
15. Kricker A, Armstrong BK, English DR, Heenan PJ. Does intermittent sun exposure cause basal cell carcinoma? a case-control study in Western Australia. Int J Cancer 1995;60:489–94.
16. Kricker A, Armstrong BK, English DR, Heenan PJ. A dose-response curve for sun exposure and basal cell carcinoma. Int J Cancer 1995;60:482–8.
17. Vajdic CM, Kricker A, Giblin M, et al. Sun exposure predicts risk of ocular melanoma in Australia. Int J Cancer 2002;101:175–82.
18. Dwyer T, Blizzard L, Gies PH, Ashbolt R, Roy C. Assessment of habitual sun exposure in adolescents via questionnaire—a comparison with objective measurement using polysulphone badges. Melanoma Res 1996;6:231–9.
19. Jones G, Blizzard C, Riley MD, Parameswaran V, Greenaway TM, Dwyer T. Vitamin D levels in prepubertal children in Southern Tasmania: prevalence and determinants. Eur J Clin Nutr 1999;53:824–9.
20. Jones G, Dwyer T, Hynes KL, Parameswaran V, Greenaway TM. Vitamin D insufficiency in adolescent males in Southern Tasmania: prevalence, determinants, and relationship to bone turnover markers. Osteoporos Int 2005;16:636–41.
21. Thieden E, Agren MS, Wulf HC. Solar UVR exposures of indoor workers in a Working and a Holiday Period assessed by personal dosimeters and sun exposure diaries. Photodermatol Photoimmunol Photomed 2001;17:249–55.
22. 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;31:439–46.
23. Last JM. A dictionary of epidemiology. New York: Oxford University Press; 1988.
24. Ponsonby AL, Lucas RM, van der Mei IA. UVR, vitamin D, and three autoimmune diseases—multiple sclerosis, type 1 diabetes, and rheumatoid arthritis. Photochem Photobiol 2005;81:1267–75.
25. Paty DW, Oger JJ, Kastrukoff LF, et al. MRI in the diagnosis of MS: a prospective study with comparison of clinical evaluation, evoked potentials, oligoclonal banding, and CT. Neurology 1988;38:180–5.
26. Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 1983;13:227–31.
27. Dwyer T, Blizzard L, Ashbolt R, Plumb J, Berwick M, Stankovich JM. Cutaneous melanin density of Caucasians measured by spectrophotometry and risk of malignant melanoma, basal cell carcinoma, and squamous cell carcinoma of the skin. Am J Epidemiol 2002;155:614–21.
28. Cohen JA. Weighted κ: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968;70:213–20.
29. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
30. Bowker AH. Bowker's test for symmetry. J Am Stat Assoc 1948;43:572–4.
31. English DR, Armstrong BK, Kricker A, Winter MG, Heenan PJ, Randell PL. Case-control study of sun exposure and squamous cell carcinoma of the skin. Int J Cancer 1998;77:347–53.
32. Feinstein AR. Indexes of contrast and quantitative significance for comparisons of two groups. Stat Med 1999;18:2557–81.
33. Blizzard L, Dwyer T, Ashbolt R. Changes in self-reported skin type associated with experience of sunburning in 14-15 year old children of northern European descent. Melanoma Res 1997;7:339–46.
34. Branstrom R, Kristjansson S, Ullen H, Brandberg Y. Stability of questionnaire items measuring behaviours, attitudes, and stages of change related to sun exposure. Melanoma Res 2002;12:513–9.
35. Berwick M, Chen YT. Reliability of reported sunburn history in a case-control study of cutaneous malignant melanoma. Am J Epidemiol 1995;141:1033–7.
36. Westerdahl J, Anderson H, Olsson H, Ingvar C. Reproducibility of a self-administered questionnaire for assessment of melanoma risk. Int J Epidemiol 1996;25:245–51.
37. Kricker A, Vajdic CM, Armstrong BK. Reliability and validity of a telephone questionnaire for estimating lifetime personal sun exposure in epidemiologic studies. Cancer Epidemiol Biomarkers Prev 2005;14:2427–32.

Validity and Reliability of Adult Recall of Past Sun Exposure in a Case-Control Study of Multiple Sclerosis

I.A.F. van der Mei, L. Blizzard, A-L. Ponsonby, et al.

Cancer Epidemiol Biomarkers Prev 2006;15:1538-1544.