

Test-Retest Reliability of the Lifestyle Questionnaire: Post-hoc Analysis of a Smoking Cessation Study

N. Mainy¹; G. Kalischnigg²; B. Zwisele²; W.T. Ng¹; P. Langer¹; M. Hankins¹; C. Haziza¹; R. Weitkunat¹

¹ PMI R&D, Philip Morris Products S.A., Quai Jeanrenaud 5, CH-2000 Neuchâtel, Switzerland

² ARGUS, Karl-Heinrich-Ulrichs-Straße 20a, DE-10785 Berlin, Germany

Introduction and Objectives

Health-related lifestyle behaviors are linked to many noncommunicable/chronic diseases. A measure of potentially confounding lifestyle behavior-related health risks was implemented in the smoking cessation study conducted by Philip Morris International (PMI) as part of the assessment of a candidate Modified Risk Tobacco Product (MRTP).

In the absence of an existing measurement instrument, the 9-item Lifestyle Questionnaire (Figure 1) was adapted from the Lifestyle Risk Scale (Alguren and Weitkunat, 2011). It measures seven health-related domains: Diet, alcohol intake, sleep deficit, physical activity, obesity, smoking, and exposure to secondhand smoke.

Special instruction: If the occurrence of the behavior is less than once (per week or days depending on the question), please answer 0.

1. How many times per week do you eat fast food? (e.g., hamburgers, hot dogs, pizza, french fries)	___
2. How many alcoholic drinks do you have a day?	___
3. How many times a week do you exercise?	___
4. What is the average duration of each exercise session? (record in minutes)	___
5. How many nights per week do you sleep less than 6 hours?	___
6. Do you live in a household with one or more smokers, other than yourself?	Yes <input type="checkbox"/> No <input type="checkbox"/>
7. On average, how many cigarettes / cigars / pipes do you currently smoke per day?	___
8. Your height:	___ cm
9. Your body weight:	___ kg

Figure 1. The Lifestyle Questionnaire.

The scores of the individual domains range from 0 to maximal 4.8, with the minimal value representing a healthy lifestyle behavior (e.g., no fast food consumption, no cigarettes smoking, more than 30 minutes of daily physical activity).

The global score is the sum of the following scores: Diet score, alcohol score, sleep score, physical activity score, smoking score, and obesity score. The global score ranges from 0 to 23.2, with higher values indicating a higher lifestyle behavior-related health risk. However, there is no scoring for the domain of exposure to secondhand smoke. The item is thus not included in the global score but gives an indication on passive smoking.

This analysis aimed at characterizing the test-retest reliability of the Lifestyle Questionnaire.

Methods

The Lifestyle Questionnaire was self-administered at screening and baseline (Visit 1 and 2, respectively) of a multicenter (42 sites), multiregional (Germany, Japan, Poland, U.K., U.S.) smoking cessation study in healthy adult smokers planning to quit smoking (registered at ClinicalTrials.gov - NCT02432729) (Figure 2). Questionnaire item #7 had a different wording (i.e. "On average, how many cigarettes per day have you smoked over the last year?"). The Screening visit to check eligibility of the subjects to participate in the study took place from 1 to 55 days prior to enrollment at the baseline visit.

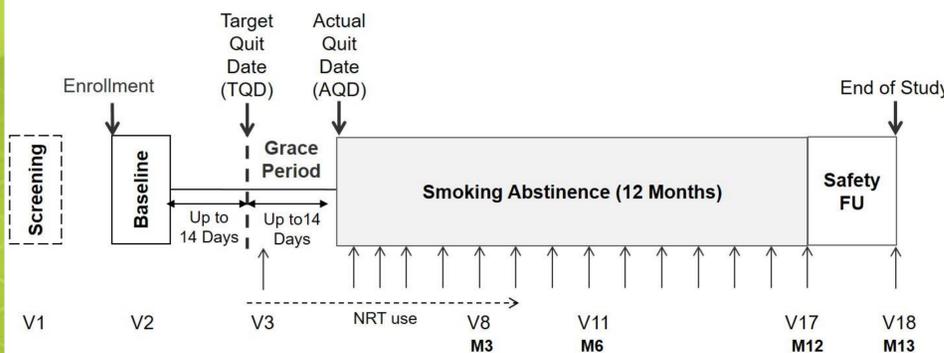


Figure 2. Smoking Cessation Study Design.

The analyses were restricted to enrolled participants with valid answers to all items of the questionnaire at both screening (Visit 1) and baseline (Visit 2).

The test-retest assessment interval between the two visits was categorized by period (Table 1).

Test-retest reliability was quantified through intra-class correlation (ICC) for the global score, and by Cohen's weighted kappa coefficient for individual questionnaire items. Stratification was conducted by test-retest assessment interval category.

Table 1. Test-Retest Assessment Interval.

Category	Period (days)
1	1 to 7
2	8 to 14
3	15 to 28
4	29 to 42
5	43 to 55

1184 subjects were enrolled in the smoking cessation study, of which 1182 provided valid answers to all items of the questionnaire at both screening and baseline visits (analysis sample). The demographic characteristics of the analysis sample are presented in Table 2. More than half of the participants (54.8%) smoked more than 15 cigarettes per day. The majority of participants (71.4%) had a normal body weight (BMI < 27.5 kg/m²).

The intraclass correlation coefficient (ICC) for the global score was 0.80, corresponding to an excellent reliability (Table 3). The ICC values varied between 0.79 and 0.90, with the highest ICC observed for the Category 1 test-retest assessment interval (i.e. < 8 days) and the lowest ICC for Category 5 (i.e. more than 42 days).

Table 3. Intraclass Correlation Coefficient for the Global Score by Test-Retest Assessment Interval.

Test-Retest Assessment Interval (days)	n	ICC [95% CI]	Classification (Cicchetti, 1994)
All	1182	0.80 [0.78; 0.82]	excellent
Cat. 1: 1 to 7	42	0.90 [0.82; 0.94]	excellent
Cat. 2: 8 to 14	310	0.80 [0.76; 0.84]	excellent
Cat. 3: 15 to 28	446	0.79 [0.75; 0.82]	excellent
Cat. 4: 29 to 42	370	0.80 [0.75; 0.83]	excellent
Cat. 5: 43 to 55	14	0.79 [0.49; 0.92]	excellent

Abbreviations: Cat.= Category; CI = Confidence Interval; ICC = Intraclass Correlation Coefficient.

Table 4 describes the global scores by test-retest assessment interval category for screening (Visit 1) and baseline (Visit 2).

Table 4. Global Score by Test-Retest Assessment Interval Category for Screening and Baseline.

Test-Retest Assessment Interval (days)	Screening (Visit 1)					Baseline (Visit 2)				
	n	Min	Mean [95% CI]	SD	Max	n	Min	Mean [95% CI]	SD	Max
All	1182	4	9.7 [9.5; 9.9]	2.71	21	1182	4	9.7 [9.5; 10.0]	2.77	23
Cat. 1: 1 to 7	42	5	9.3 [8.4; 10.2]	2.80	18	42	5	9.2 [8.3; 10.1]	2.71	16
Cat. 2: 8 to 14	310	4	9.8 [9.5; 10.1]	2.63	18	310	4	9.8 [9.4; 10.1]	2.66	18
Cat. 3: 15 to 28	446	4	9.7 [9.4; 10.0]	2.77	21	446	4	9.8 [9.5; 10.1]	2.75	23
Cat. 4: 29 to 42	370	4	9.6 [9.3; 9.9]	2.68	21	370	4	9.7 [9.4; 10.1]	2.88	20
Cat. 5: 43 to 55	14	6	10.3 [8.5; 12.0]	2.94	15	14	5	10.0 [8.1; 12.0]	3.33	17

Abbreviations: Cat.= Category; CI = Confidence Interval.

The inter-specific percent agreement at screening and baseline varied between 58.0% and 94.3%, with the lowest agreement for Item #1 Diet, and the highest agreement for Item #8 BMI (Table 5). No significant differences were observed in the percent agreement among test-retest assessment interval categories.

Table 5. Agreement between Screening and Baseline Visits by Questionnaire Item.

Questionnaire Item	Agreement % [95% CI]
Diet: Fast food per week (#1)	57.95 [55.08; 60.79]
Alcohol Intake: Alcohol per day (#2)	79.44 [77.02; 81.71]
Physical activity: frequency (#3)	69.97 [67.26; 72.57]
Physical activity: duration (#4)	72.42 [69.78; 74.95]
Sleep deficit: Nights per week with sleep under 6 h (#5)	58.71 [55.85; 61.54]
Exposure to secondhand smoke (#6)	93.32 [91.74; 94.67]
Smoking: Cigarettes per day (#7)	86.97 [84.92; 88.84]
Body Mass Index (BMI) (#8 & #9)	94.33 [92.86; 95.58]

Abbreviations: CI = Confidence Interval.

Cohen's weighted Kappa varied between 0.70 and 0.96, with substantial agreement for Item #1 Diet, Item #2 Alcohol intake, Items #3 and #4 on Physical activity (frequency and duration, respectively), Item #5 Sleep deficit, and a high agreement for Item #6 Secondhand smoke, Item #7 Smoking, and Item #8 (Table 6).

Results

Table 2. Demographic Characteristics.

Variable	Category	All (N=1182)	
		n	%
Age Group	30-39	424	35.87
	40-49	414	35.03
	50-65	344	29.10
Gender	Female	580	49.07
	Male	602	50.93
Country	Germany	104	8.80
	Japan	354	29.95
	Poland	281	23.77
	UK	103	8.71
	US	340	28.77

Table 6. Weighted Kappa Coefficients.

Questionnaire Item	Weighted Kappa coefficient [95% CI]	Classification (Landis and Koch, 1977)
Diet: Fast food per week (#1)	0.72 [0.68; 0.76]	substantial
Alcohol Intake: Alcohol per day (#2)	0.73 [0.69; 0.78]	substantial
Physical activity: frequency (#3)	0.79 [0.76; 0.83]	substantial
Physical activity: duration (#4)	0.77 [0.74; 0.81]	substantial
Sleep deficit: Nights per week with sleep under 6 h (#5)	0.70 [0.66; 0.74]	substantial
Exposure to secondhand smoke (#6)	0.86* [0.82; 0.89]	almost perfect
Smoking: Cigarettes per day (#7)	0.85 [0.82; 0.88]	almost perfect
Body Mass Index (BMI) (#8 & #9)	0.96 [0.95; 0.97]	almost perfect

Abbreviation: CI = Confidence Interval.

* Unweighted Kappa coefficient

Figure 3 depicts weighted Kappa values by questionnaire item and test-retest assessment interval.

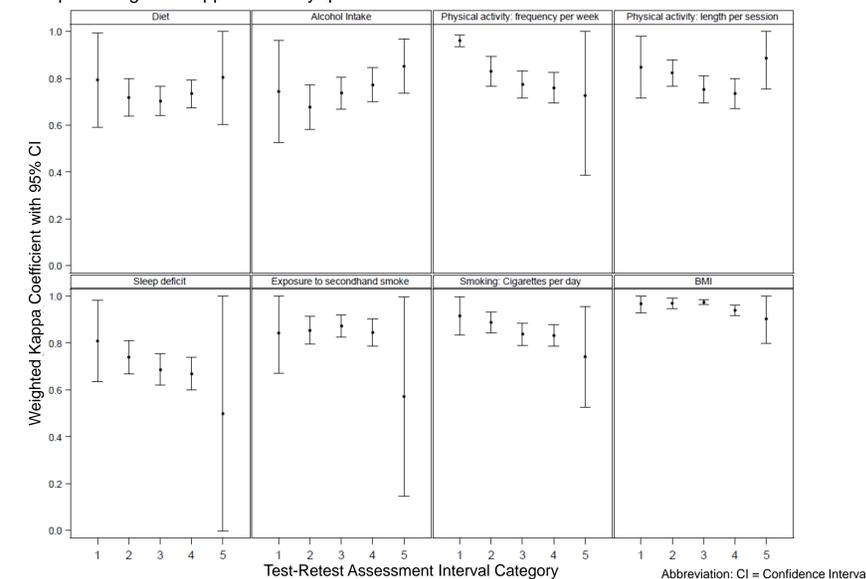


Figure 3. Weighted Kappa Coefficient per Questionnaire Item and Test-Retest Period.

Discussion and Conclusions

The ICC values showed excellent agreement for all test-retest assessment intervals with the precision varying through the considered test-retest assessment interval categories due to the different group sizes. The global score showed an excellent stability of the responses (ICC = 0.80).

The answers to the individual questionnaire items were also stable across the two visits.

The findings regarding the reliability of self-report on smoking (i.e. number of cigarettes per day) are in line with previously published results (Sponsiello-Wang et al., 2017).

Measuring health-related lifestyles as potential confounding factors is key to many epidemiological research areas, as well as for the assessment of tobacco- and nicotine-containing products, such as Modified Risk Tobacco Products (MRTP). The Lifestyle Questionnaire was found to provide a short and simple yet reliable instrument to characterize potential health-related behaviors in clinical and epidemiological studies.

The Lifestyle Questionnaire in English for the U.S., together with existing linguistically validated translations to six languages (English for the U.K., German for Germany, Japanese for Japan, Polish for Poland, Spanish for the U.S., and Swedish for Sweden) is available on PROQOLID™.

References:

- Alguren B, Weitkunat R. Perception and prevalence of behavioral risk factors: the lifestyle risk scale (LRS). Open Journal of Preventive Medicine. 2011;1(3):143-153.
- Sponsiello-Wang Z, de La Bourdonnaye G, David M, Lüdiche F, Weitkunat R. Accuracy of the smoking questionnaire. Contributions to Tobacco Research. 2017;27(8).
- Landis J.R., Koch G.G. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174.
- Cicchetti D.V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 1994; 6(4), 284-290.