

# Assessment of Cigarette Smoking in Epidemiologic Studies \*

by

Rolf Weitkunat<sup>1</sup>, Christopher R.E. Coggins<sup>2</sup>, Zheng Sponsiello-Wang<sup>1</sup>, Gerd Kallischnigg<sup>3</sup>, and Ruth Dempsey<sup>1</sup>

<sup>1</sup> PMI Research & Development, Philip Morris Products SA, Neuchâtel, Switzerland

<sup>2</sup> Carson Watts Consulting, 1266 Carson Watts Road, King, NC, USA

<sup>3</sup> ARGUS - Statistics and Information Systems in Environment and Public Health, Einemstr. 20a, 10785 Berlin, Germany

## SUMMARY

The assessment of cumulative exposure based on collecting information on the history of active cigarette smoking has been and is being undertaken in a variety of ways. While a very detailed assessment may be required for studies with a focus on particular aspects of smoking behavior and history, comparability of measurements and results across studies remains a primary concern. Addressing the problem of heterogeneity of exposure assessment across studies can be achieved by a core set of questions that cover the major dimensions of cigarette smoking, and yet comply with current criteria used for defining smoking history and status. In studies where no very high level of exposure assessment is required or where smoking is not the major subject of investigation, a practical standardized core set of questions appears to be of considerable value, in particular with regard to making results more comparable across studies. [Beitr. Tabakforsch. Int. 25 (2013) 638–648]

## KEY WORDS

Exposure assessment, active smoking, smoking questionnaire, standardized measurement

## ZUSAMMENFASSUNG

Die Messung der kumulativen Exposition durch Fragen zum biographischen Rauchverhalten erfolgt bis heute nicht auf einheitliche Weise. Auch wenn einzelne Studien zu ganz bestimmten Aspekten des gegenwärtigen und früheren Rauchverhaltens eine sehr differenzierte Datenerhebung erforderlich machen mögen, bleibt doch die eingeschränkte

Vergleichbarkeit von Erfassungsmethoden und Resultaten über Studien hinweg ein Hauptproblem. Durch einen Satz zentraler Fragen, die die Hauptdimensionen des Zigarettenrauchens abdecken und die mit vorliegenden Definitionen des Rauchverhaltens konsistent sind, kann das Problem der Heterogenität der Expositionsmessung entschärft werden. Für Studien, die keine hochdetaillierte Expositionsmessung erfordern oder in denen Rauchen nicht die Hauptfragestellung darstellt, ist ein standardisierter Fragensatz, insbesondere im Hinblick auf die Vergleichbarkeit verschiedener Studien, von offensichtlichem Vorteil. [Beitr. Tabakforsch. Int. 25 (2013) 638–648]

## RESUME

L'évaluation de l'exposition cumulée à la cigarette chez le fumeur sur la base d'informations collectées sur son passé de fumeur a été et est toujours menée de manières très variées. Alors qu'une évaluation très précise pourrait être nécessaire dans les études focalisant sur certains aspects du type de fumeur et son passé, la comparabilité des mesures et des résultats entre études reste la principale préoccupation. Le problème d'hétérogénéité des mesures entre les études peut être résolu par l'utilisation d'un ensemble de questions couvrant les dimensions principales relatives au comportement du fumeur, qui de plus répondrait aux critères actuels utilisés pour définir le type de fumeur ainsi que son passé. Dans les études qui ne requièrent pas une évaluation de l'exposition très fine ou pour lesquelles le tabac n'est pas le sujet d'étude majeur, l'utilisation d'un tel ensemble de questions standardisées peut s'avérer très judicieuse, notamment afin de pouvoir mieux comparer les résultats entre études. [Beitr. Tabakforsch. Int. 25 (2013) 638–648]

## INTRODUCTION

Over the past 60 years cigarette smoking has been one of the most investigated risk factors in epidemiology (1, 2). From an epidemiological point of view it is a highly prevalent exposure causally associated with a variety of diseases. Therefore, even in studies not focused on smoking, its role as a potential confounder must usually be carefully considered. Astonishingly, the definition of active smoking exposure measurements has not been fully standardized. In 1950 DOLL and HILL (1) pointed out the necessity to define what was meant by “smoker”, and considered having smoked every day for at least one year as defining ever (i.e., lifetime) smoking. In a review of 18 major US surveys dealing with smoking, MASSEY *et al.* (3) have reported eight different definitions of “smoker”, let alone further details of smoking behavior. The authors pointed out that the National Clearinghouse on Smoking and Health had already in 1955 provided a definition, based on a lifetime consumption of 100 or more cigarettes and currently smoking, which was not adopted by all surveys, even two to three decades later. As of 1965, the 100 cigarette question was used in the US National Health Interview (NHIS) (4) to define ever smoking and subsequently became a *de facto* criterion for adult lifetime, former, and current smoking in health studies modeled after the NHIS (5). It should be noted that definitions differ for adolescents, and smoking even once or twice or even one or two puffs may qualify for ever smoking, depending on the definition (6). While from 1965–1991, simply reporting current smoking was the NHIS required additional criterion for ever smokers to be current smokers, as of 1992 current smoking at least on some days was required (6). Also, as of 1992, former smokers (who previously were defined as ever smokers who had to report that they no longer smoked) have been defined as ever smokers that now did not smoke, neither daily nor on some days. Other than the US NHIS, the European Core Health Interview Survey (ECHIS) questionnaire (7) is recommended as a European reference standard for national health surveys. It does not contain a question indicating consumption of 100 cigarettes and only proposes a set of five questions on active smoking. By measuring current non-smoking, current occasional smoking (not being further defined), current daily smoking (including type and number of products smoked by current daily smokers) as well as duration (in years) of ever “daily, or almost daily” smoking, little more than prevalence estimates can be derived. The European standard renders comparisons to NHIS-based data difficult and also falls short for providing information on various definitions of tobacco use as proposed by the WORLD HEALTH ORGANIZATION (WHO) in 1998 (8). These are based on twelve categories, including never, non, ever, current, and occasional smoking. While the WHO definitions have been widely used and allow for a classification of smoking behavior and history, they lack specification, e.g., on the duration of smoking or cessation, on smoking intensity, or on smoking patterns, e.g., with regard to occasional smoking. Overall, the lack of consistency of smoking exposure measurement and definition is not yet fully resolved and was, by MASSEY *et al.* (3), considered a “primary concern when using survey data to describe

*trends...*” (p. 433). Similarly, the INTERNATIONAL ASSOCIATION FOR RESEARCH ON CANCER (IARC) pointed out in 2007 (9) that “*proper classification of smoking behavior is critically important*” in the context of evaluating the health consequences of smoking and cessation, and examples of varying definitions of smoking and ex-smoking across studies and countries are given. DELNEVO and BAUER (10) have recently reviewed ten major US surveys which collected data on tobacco use. The 100 cigarette question was not included in two core questionnaire modules, and three different surveys did not collect information on daily or occasional smoking, the number of cigarettes smoked per day, or quit attempts. Age of first use of cigarettes, duration of abstinence, and information on brands (which would provide a basis for information on “tar” and nicotine levels) was not collected in three, four, and six surveys, respectively. In nine surveys either no or incomplete information was collected on other tobacco products (including cigars, pipes, roll-your-own cigarettes, and smokeless tobacco). The authors of this review concluded that “*a systematic understanding of the tobacco host remains unavailable due to gaps in survey content...*” (p. S22).

Conducting an internet search for the phrase “current smoking was defined”, restricted to the most recent year, quickly demonstrates the persistence of a striking variability of smoking definitions and measurements across epidemiological and clinical studies. This is also reflected in a recent comparison of smoking prevalence estimates from the European Commission Eurobarometer survey of 27 European Union member states versus estimates provided by national survey data (11). The authors compared smoking-related questions of the Eurobarometer and of national surveys using, among other sources, the European Health Interview & Health Examination Surveys Database (<https://hishes.iph.fgov.be>). National prevalence estimates differed by up to 13 percentage points from the country-specific Eurobarometer estimates. The authors stated that “*the questions on smoking used in national survey varied markedly in phrasing...*” (p. 2) and that “*harmonisation of methodology, or at least the inclusion of a basic set of common questions, would therefore be a major step towards the generation of data suitable to compare smoking prevalence in European countries over place and time*” (p. 8). As noted in an analysis of even a single (but quite critical) variable related to “smoking status” (100 cigarettes ever smoked), “*there are many differences among definitions of current smoking, including whether occasional or non-daily smokers are included and whether an explicit time frame is included. Definitions of lifetime smoking vary as well, and include everything from a single puff through to daily smoking for a year*” (5). Complementary to the problems of defining what a “smoker” is, there is also confusion over the meaning of “non-smoker” (12). While linguistically English (and German) suggests that a non-smoker is someone who does not currently smoke (never and ex combined), the term is often used to imply never smoking (when this should strictly be lifelong nonsmoking).

Clearly, for most research objectives, more information than “smoking status” is needed. As noted by DIETRICH and HOFFMANN (13) “*the use of a dichotomous variable for*

*'ever-smoking' will result in an estimate that is easy to interpret, but most of the variability in smoking history will not be captured*". While the determination of the "smoking status" of study participants (usually categorized as current, ex-, or never-smokers) provides a rather general level of exposure assessment, the dimensions of intensity, age at initiation, duration of smoking, duration of smoking cessation, pattern of smoking and type of cigarette (in terms of tobacco type, "tar" level, ingredients, design, type of filter, cf. (14–17)) are not considered in such a simple classification. With regard to pattern, it was observed that occasional (non-daily) smoking can be a stable pattern for long periods (18) and that the prevalence of occasional smoking has increased in the US (19). While non-daily smoking can occur on a more or less random basis, some intermittent smokers (chippers) may smoke only on certain days of the week or in certain contexts or settings (20). The importance of exposure assessment beyond classification of smoking status has been demonstrated in various respects. More or less complex questionnaires have been useful to quantify the most important factors driving observed dose-response relationships. For example, smoking duration has been found to be more important than smoking intensity (in terms of cigarettes per day) in causing lung cancer (21–25). Also, temporal changes of smoking behavior (smoking trajectories) are of relevance, but are not considered by a simple classification of smoking status. A sufficiently detailed assessment of active smoking related exposure is also needed for determining important derived variables. An often used integrative metric to quantify cumulative smoking exposure ("smoking history") is "pack-years", sometimes given as "cigarette years", being defined as the product of number of cigarettes smoked per day (expressed in units of packs of 20 cigarettes smoked) and of years of smoking (26).

## OBJECTIVES

In this paper a simple Smoking Questionnaire (SQ) is proposed which can be easily applied in many epidemiologic research contexts and provides a range of measures of smoking exposure in a simple and standardized manner. Devising the SQ as a basic minimum standard for measuring individual exposure related to active cigarette smoking behavior in epidemiologic studies was guided by a set of attributes which are listed in the approximate order of importance:

- useful, i.e., the SQ should be suitable to specifically assess exposure to active cigarette smoking for a wide range of research objectives and study designs, implying it is informative (nothing essential is missing), concise (assessment is efficient, redundancy is avoided), and practical (simple, clear, swift), thus maximizing robustness and reliability;
- sufficient, i.e., the SQ should completely collect all relevant information at adequate levels of granularity, including exposure type and dosing in time, as well as information on non-cigarette smoking, to allow for controlling bias and confounding;
- coherent, i.e., the SQ should be based on available knowledge on relevant aspects of the exposure, thus maximizing logical and content validity;
- consistent, i.e., the SQ should be backward compatible with previous questionnaires and definitions, including response metrics and categories, thus maximizing criterion validity and minimizing systematic errors (bias) as well as supporting comparability with previous research.

Questions related to more complex research objectives are not in the scope of the present proposal but, as will be demonstrated below, the information collected with the SQ allows for deriving of a wide range of variables.

## QUESTIONNAIRE

The proposed SQ (version 1.0, cf. Table 1) is intended for use in studies which do not require a very detailed assessment of current or lifetime cigarette smoking. Its focus is on self-reported current and past cigarette smoking behavior, allowing derivation of many of the major definitions that have been previously utilized, as will be demonstrated below. History of use of other smoking products is quantified, in addition to cigarette smoking, in order to allow identification of different sources of tobacco smoke exposure.

Although the order of the SQ questions should be maintained, the instrument is meant as a set of questions to be included in a larger study-specific questionnaire or case report form. It appears appropriate for paper-based or electronic self-administration, as well as for use in interviews, either face-to-face or through telephone. Translation into other languages should be straightforward. It should be noted that the date of the assessment, the date of birth and/or the age of the respondent are assumed to be acquired separately in addition to the data captured with the SQ. Rather than attempting to obtain all relevant variables analysis-ready directly from the respondents, the SQ focuses on simplicity of use. It places the burden of effort for deriving the relevant information on the post-processing and analysis of the acquired data. The advantage is that all data required for further processing can be collected in a simple and brief manner, naturally at the expense of requiring further processing of the captured data before rendering most of its utility.

The first of the eight questions allows for a broad classification of current cigarette smoking behavior (daily, occasional, ex, non). The ordering of the response categories is guided by the rationale that by reversing the order, respondents might be tempted to answering "non-smoker" before noticing (and then possibly not even reading) the option "ex-smoker". Questions 2 and 3 collect information on the 100 cigarettes criterion and on regular ever smoking (defined as smoking at least one cigarette per day). While the information explicitly collected by questions 1 and 3 can technically be derived from the responses on questions 8 and 4, respectively, their pivotal importance in defining smoking status and the small effort required to answer them seems to outweigh this redundancy. Question 4 captures

**Table 1. Smoking questionnaire.** Age and sex of the respondent are required for a comprehensive assessment, but are not contained in the questionnaire. It is assumed that this information is obtained separately.

No.	Question	
1	What is your current cigarette smoking behavior (including hand-rolled cigarettes)?	
A	Daily smoker (at least one cigarette per day, disregarding religious fasting)	<input type="checkbox"/>
B	Occasional smoker (less than one cigarette per day)	<input type="checkbox"/>
C	Ex-smoker of cigarettes	<input type="checkbox"/>
D	Non-smoker of cigarettes	<input type="checkbox"/>
2	Have you ever smoked 100 cigarettes or more in your life?	
A	Yes	<input type="checkbox"/>
B	No	<input type="checkbox"/>
3	Did you ever smoke cigarettes regularly, i.e., at least one cigarette per day?	
A	Yes	<input type="checkbox"/>
B	No	<input type="checkbox"/>
4	If you ever smoked cigarettes regularly: At what age did you start to smoke regularly?	___ years
5	If you are an ex-smoker of cigarettes: For how long have you quit now?	___ years + ___ months + ___ days
6	If you ever quit regular cigarette smoking: For how long did you quit altogether? Please add together any separate periods of quitting.	___ years + ___ months + ___ days
7	What brand of cigarettes / hand-rolled tobacco did you predominantly smoke in the last 12 months of smoking?	Name of brand: _____
8	On average, how many cigarettes / cigars / pipes do / did you smoke per day?	

  

		Manufactured cigarettes	Hand-rolled cigarettes	Cigars	Pipes
A	Currently (last 3 months)	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day
B	1 year ago	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day
C	5 years ago	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day
D	10 years ago	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day
E	15 years ago	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day
F	20 years ago	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day
G	> 20 years ago	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day	<input type="checkbox"/> none <input type="checkbox"/> less than 1 per day ___ per day

age of initiation and questions 5 and 6 current as well as total quitting duration, respectively. The response format for quitting durations is a combination of years, months, and days, thus allowing for capturing information at different levels of detail, according the individual accuracy of recall. The brand of cigarettes predominantly smoked in the last 12 months of smoking is captured in question 7, which allows for linking additional information to the data set, based on available background data, e.g., on nicotine and “tar” levels, tobacco type, pack size, filter type, and

cigarette length. Question 8 captures the tobacco smoking history. Daily numbers of manufactured as well as of hand-rolled cigarettes, cigars, and pipes are acquired, separately for seven time periods. For completeness and consistency of the collection of information on smoking intensity in terms of daily product use frequency, not smoking a particular product type at a certain time is to be indicated, as well as smoking of less than one unit per day, thus fully covering the spectrum of possible smoking behaviors. These two response categories also help to avoid perplexing

younger respondents by asking them about a past period where they could not have possibly smoked. The temporal resolution of question 8 is largest for the most recent periods (last 3 and 12 months), and by 5 year-intervals thereafter. The last period captures smoking intensity of more than 20 years ago.

While some questionnaires organize the acquisition of consumption levels by age, this does not avoid respondents having to recall behaviors that they had practiced many years ago. Asking for consumption at certain age levels (and abstaining from routing) implies that the inquiry needs to commence at a relatively early age, presumably 15 years, in order to obtain minimum-length retrospective time intervals also for younger participants. Since all age groups need to be foreseen, many age levels not applicable to younger participants would need to be included, whereas older participants would then be charged with historic behaviors beyond the range of recall. While considering the possible age range would already lead to a longer question 8 than the one actually proposed, current consumption (covering the last three and twelve months as in the present version) would still have to be requested in addition. This would not only further lengthen the questionnaire, but also lead to a mixture of time reference systems (age and chronological time), and would increase the level of complexity of the question and thus the cognitive burden placed on the respondents.

No routing based on linking responses to filter functions that would allow skipping certain parts was built into the SQ, even though filters could reduce the time required to fill in the questionnaire, in particular for never-smokers. Since, however, there appears to be confusion about non and never smoking it appeared to be appropriate not to have a never-smoker category in question 1, but rather to derive this status from the information acquired uniformly from all (including never smoking) respondents in the subsequent questions. Another possibility for routing would be to request detailed information only from self-assessed ever regular smokers (question 3). This (and any) routing would, however, impose restrictions on the possibilities to analyze the data, as it would imply conditioning subsequent data on previous responses to filter questions, thus imposing some type of a-priori classification and definitions, with consequences that are hard to foresee for all possible research contexts. In contrast, abstaining from any routing, and instead focusing on brevity and simplifying the cognitive requirements for filling in the questionnaire, allows for flexibility in definitions used for classifying study participants during the data analysis, e.g., with regard to never smoking to be or not to be based on the 100 cigarettes criterion. It should also be considered that filters have the potential to introduce bias to a questionnaire, based on respondent's prospects of saving time to fill in subsequent questions conditional to a certain response, e.g., checking the non-smoking category in question one. Avoiding routing also considerably alleviates constraints to adhere to a strictly logical architecture, regarding both the sequence and formulation of questions and response categories, and, most importantly, allows for a respondent-oriented rather than analysis-logical questionnaire design with a focus on ease of use.

## POST-PROCESSING

While the information gathered with the SQ may not be directly suitable for analysis for all purposes, it supports indirect derivation of many variables and indicators. To demonstrate the respective utility of the SQ, Table 2 contains the twelve 1998 WHO basic definitions of tobacco use (8). According to these, the main criteria for being classified as a "smoker" are current daily smoking or smoking occasionally and having "*ever smoked at least 100 cigarettes (or the equivalent amount of tobacco) in their lifetime.*" The WHO categories of tobacco use, in addition to "smoker", are "daily smoker", "occasional smoker", "reducer", "continued occasional smoker", "experimenter", "non-smoker", "ex-smoker", "never-smoker", "ex-occasional smoker", "ever smoker", and "ever daily smoker". By combining the data captured with the SQ it is possible to classify responders according to these categories, both with regard to cigarette smoking as well as to smoking in general. In Table 2, the two columns to the right contain the logical conditions required to achieve the classification of SQ data according to the WHO definitions. They can be easily translated into data management statements to computerize the classification of a study sample.

For example, to qualify as a daily cigarette smoker, response option A of question 1 (at least one cigarette per day) must have been checked ( $1A$ ), and a daily smoker in general is identified by either daily cigarette smoking or current daily smoking of any of the products covered in question 8 ( $1A \cup CP_A \geq 1$ ). Alternatively, to qualify as a never cigarette smoker, response option D of question 1 (non-smoker of cigarettes) must have been checked, as well as option B of question 2 (not having smoked 100 cigarettes altogether) and option B of question 3 (never having smoked cigarettes regularly) ( $1D \cap 2B \cap 3B$ ). Additional requirements for never smoking in general are current non-smoking of any product covered in question 8, as well as never daily smoking, indicated by the sum of any products smoked across all periods covered in question 8 being less than one ( $1D \cap 2B \cap 3B \cap CP_A = 0 \cap \sum_{i=A,G} CP_i < 1$ ).

As can be seen, all WHO basic definitions regarding cigarette smoking can be addressed unequivocally. For the sake of brevity, smoking behavior regarding cigars and pipes is not measured at the same level of detail with the SQ, as it does, for example, not contain questions on daily and occasional pipe and cigar smoking. As a consequence, referring to smoking in general, not all categories of the WHO basic definitions can be categorized in a completely definite way. This impacts the definition of smoker, daily smoker, and occasional smoker, as current smoking of pipes and cigars is only assessed on the basis of average daily numbers in question 8.

Of note, the WHO definitions are rather basic and do not give, for example, any indication of how long smokers had to have not smoked before they would be considered as ex-smokers. To avoid very recent quitters classifying themselves as ex-smokers, a minimal period of quitting (for example three months) should probably be considered as a criterion to qualify as an ex-smoker. The SQ uses in question 1 the unspecific response category of the WHO definition, but question 5 provides detailed quantification of the current quitting period in ex-smokers, therefore

**Table 2. WHO 1998 basic definitions of tobacco use.** Correspondence of the definitions with the response patterns of the smoking questionnaire is contained in the columns to the right. Correspondence with regard to definitions of smoking in general is largely based on the data captured in question 8.  $CP_i$  indicates the sum of the average daily numbers of cigars and pipes smoked in period  $i$  (ranging from A to G in question 8). If other non-cigarette smokes are to be assessed by additional columns in question 8, then  $CP_i$  must be adjusted accordingly. The operators  $\cap$  and  $\cup$ , used with or without index  $i$  to indicate repetition across periods, indicate logical conjunction (logical "and") and disjunction (logical "or"), respectively, of the connected response categories. The  $\sum_i$  operator indicates summation across periods. Negation is indicated by the  $\neg$  operator.

WHO Definition		Corresponding smoking questionnaire response patterns regarding	
		Cigarette smoking	Smoking in general
A	A <i>smoker</i> is someone who, at the time of the survey, smokes any tobacco product either daily or occasionally.	$1B \cup 1A$	$1B \cup 1A \cup CP_A > 0$
A1	A <i>daily smoker</i> is someone who smokes any tobacco product at least once a day (with the exception that people who smoke every day, but not on days of religious fasting, are still classified as daily smokers).	1A	$1A \cup CP_A \geq 1$
A2	An <i>occasional smoker</i> is someone who smokes, but not every day. Occasional smokers include:	1B	$[1B \cup (0 < CP_A < 1)] \cap \neg (1A \cup CP_A \geq 1)$
A2i)	<i>Reducers</i> - people who used to smoke daily but now do not smoke every day.	$1B \cap 3A$	$[1B \cup (0 < CP_A < 1)] \cap (3A \cup \bigcup_{i=B,G} CP_i \geq 1) \cap \neg (1A \cup CP_A \geq 1)$
A2ii)	<i>Continuing occasional</i> - people who have never smoked daily, but who have smoked 100 or more cigarettes (or the equivalent amount of tobacco) and now smoke occasionally.	$1B \cap 2A \cap 3B$	$[1B \cup (0 < CP_A < 1)] \cap (\sum_{i=A,G} CP_i \geq 1) \cap \neg (1A \cup 3A \cup \bigcup_{i=A,G} CP_i \geq 1)$
A2iii)	<i>Experimenters</i> - people who have smoked less than 100 cigarettes (or the equivalent amount of tobacco) and now smoke occasionally.	$1B \cap 2B$	$[1B \cup (0 < CP_A < 1)] \cap (2B \cap \sum_{i=A,G} CP_i < 1) \cap \neg (1A \cup CP_A \geq 1)$
B	A <i>non-smoker</i> is someone who, at the time of the survey, does not smoke at all. Non-smokers can be divided into three categories:	$1D \cup 1C$	$(1D \cup 1C) \cap CP_A = 0$
B1	<i>Ex-smokers</i> are people who were formerly daily smokers but currently do not smoke at all.	$1C \cap 3A$	$(1C \cup CP_A = 0) \cap (3A \cup \bigcup_{i=B,G} CP_i \geq 1) \cap \neg (1B \cup 1A \cup CP_A > 0)$
B2	<i>Never-smokers</i> are those who have never smoked at all or have never been daily smokers and have smoked less than 100 cigarettes (or the equivalent amount of tobacco) in their lifetime.	$1D \cap 2B \cap 3B$	$1D \cap 2B \cap 3B \cap CP_A = 0 \cap \sum_{i=A,G} CP_i < 1$
B3	<i>Ex-occasional smokers</i> are those who were formerly occasional, but never daily, smokers and who smoked 100 or more cigarettes (or the equivalent amount of tobacco) in their lifetime.	$1C \cap 2A \cap 3B$	$[(1D \cup 1C) \cap CP_A = 0] \cap (2A \cup \sum_{i=B,G} CP_i > 1) \cap (3B \cup \neg \bigcup_{i=B,G} CP_i \geq 1) \cap \neg (3A \cup \bigcup_{i=B,G} CP_i \geq 1)$
C	These definitions can be used to classify the population according to their lifetime smoking status. In particular: <i>Ever smokers</i> are defined as those who have ever smoked at least 100 cigarettes (or the equivalent amount of tobacco) in their lifetime. A specific subcategory of interest is those who have smoked, or now smoke, every day:	2A	$2A \cup \sum_{i=A,G} CP_i \geq 1$
C1	<i>Ever daily smokers</i> are defined as persons who are currently daily smokers, reducers, or ex-smokers.	$2A \cap (1A \cup 3A)$ (alternatively: disjunction of definitions A1, A2i, and B1)	$(1A \cup 3A \cup \bigcup_{i=A,G} CP_i \geq 1) \cap (2A \cup \sum_{i=A,G} CP_i \geq 1)$

allowing for a flexible re-classification during data analysis.

Use of multiple types of tobacco products can be considered, based on the quantitative information collected in question 8 of the SQ. Categories like "almost exclusive use", "predominant use", "equal (or average) use", "occasional use", "rare use", etc. can be derived by specifying ranges of proportionate use of particular tobacco products. Alternatively, weight-based tobacco equivalents can be determined for different tobacco types. For example, it has been suggested that the tobacco equivalents for cigarettes should be 1 g, for cigarillos 2 g, and for pipes and cigars 3 g (27). It should be noted, however, that there are diffi-

culties with these measures in that 1g appears to be too high based on the tobacco weight distribution of manufactured cigarettes (28), cigars may vary considerably in format, and there may be period and country specific differences in tobacco product weights. The SQ data can also be used to calculate aggregated measures of lifetime smoking exposure beyond pack years, incorporating intensity, duration, and the time since cessation (13, 29, 30). Furthermore, secular categories of smoking intensity levels (e.g., increasing, reducing, constant) can be derived from the data collected with question 8.

By interpolation and extrapolation (to the time of initiation), estimating lifetime consumption becomes possible,

including the calculation of pack years. Also, the different specific periods in the past (as specified in question 8) can be converted to years of age at that time, by subtracting them from the current age of the respondent. For any desired age (e.g., 20, 30, 40 etc. years of age), consumption levels can then be estimated from the available data of before and after that age through simple linear interpolation or by using methods to account for the overall nonlinear shape of the consumption trajectory (e.g., polynomial or spline interpolation).

## DISCUSSION

The availability of a standardized basic smoking questionnaire can support the quality of smoking exposure assessment and the comparability of results across studies (31). The major goal of this work was to propose a minimum standard, useful, sufficient, coherent, and consistent set of questions to assess active cigarette smoking, including cumulative smoking history. The proposed SQ is specifically designed for use in studies which do not require a very high level of detail of active cigarette smoking. It is focusing on brevity and simplicity and shifts the load for providing analysis-ready variables from the respondent to the stage of data post-processing. In addition to suggesting a practical standard for future research, the SQ provides a framework for analyzing or re-analyzing existing data captured with other questionnaires, to increase the comparability of results.

The time requirement to fill in the first seven of the eight SQ questions is likely to take less than a minute. To fill in question 8, where smoking history is captured, can require more effort, as it covers various periods of the present and past as well as four different types of tobacco smoking products; for each applicable combination (i.e., when the respondent smoked a certain product type at a certain time), a quantification of the number of units smoked is required. However, only respondents who are about 40 or more years of age and who had started smoking all four product types more than 20 years ago must provide numbers to all cells of question 8, whereas the majority of smokers most likely will have to provide quantities only for some periods and products. It seems that the majority of smokers answering question 8 might need to spend again less than a minute on this part of the SQ, so that the overall time required from an “average smoker” should not exceed two minutes for the whole questionnaire.

The SQ is strictly focused on current and former active cigarette smoking behaviour and does not provide all the details that can be captured by more comprehensive questionnaires. Some instruments cover in detail very specific aspects related to smoking such as exposure to second-hand smoke, awareness or response to advertisements or anti-smoking information, prices paid, health-related lifestyle variables aside of (but potentially related to) smoking, craving and withdrawal symptoms, or craving (e.g., Lifetime Tobacco Use Questionnaire LTUQ (32), key questions from the Global Adult Tobacco Survey GATS (33), Lifetime Interview on Smoking Trajectories LIST (34) or Fagerström Test for Nicotine Dependence FTND (35)). While being reasonably short (albeit not as short as,

for example, the ECHIS, which has only five questions on active smoking), the SQ facilitates the derivation of a wealth of information, including all WHO smoking status classes, rather than only providing information on a subset thereof (e.g., current and ever smoking, to allow for simple prevalence estimates). By defining a simple standard of measuring active smoking behavior, the SQ might contribute to bringing some clarity to smokers themselves, regarding their smoking status; it has been reported that many smokers appear to be insecure about whether or not to consider themselves as smokers (36), which can in turn contribute to misclassification problems in studies using only a very general classification of smoking status (cf. (37)).

The SQ is not conceived as a stand-alone instrument that must, under all circumstances, be used as is. If the SQ is extended, however, it is important to ensure that the questions and categories are maintained as proposed, otherwise there is a risk that one of the major advantages of the SQ could be lost, i.e., standardization and comparability across studies. If possible, additional questions should therefore be added to the end of the SQ to avoid response bias due to altered question sequence. For example, in order to address smoking in the past 30 days (or other recent periods), one or more additional questions to measure on how many days cigarettes were smoked during that period would allow for assessing current daily and non-daily smoking (e.g., (36)), with a specification of “current” that is also used in the NHIS.

To further increase the temporal and behavior-related resolution, it is possible to add to question 8 additional rows referring to other periods in the past, as well as columns for other than the four currently assessed tobacco products (possibly including noncombustible, smokeless, or dissolvable products, waterpipes, e-cigarettes, reduced exposure products, modified risk tobacco products etc.) - to more comprehensively assess dual and polytobacco use - as well as nicotine replacement therapies. To quantify consumption levels at a higher resolution than currently foreseen for the lower end of the distribution, the number of units smoked per week could be acquired in question 8, in addition to the current per-day quantification, although it seems doubtful whether recall at such granularity would be reliable for various periods in the past. Examples of more particular extensions of the current SQ include adding questions on daily and occasional pipe and cigar smoking (in analogy to the assessment of cigarette smoking) or the collection of consumption intensities stratified by day of the week, time of the day, or social context.

In its present form, the SQ can be used in cross-sectional, case-control, as well as in prospective studies (with cigarette smoking assessment at time of study onset). For repeated deployment, questions 4 and 6 could be removed and the timelines in question 8 adjusted according to the planned follow-up intervals. When applied in appropriate study designs, in addition to prevalence rates and derived rate ratios (e.g., the “quit ratio” of former to ever smokers), the SQ can provide information required for incidence rate calculations (e.g., rates of cessation, initiation, relapse, product switching, intensity changes etc.). When extending the SQ, it is advisable to carefully consider which derived variables are actually needed and how these should be used

in the analysis. Issues as to which variables to combine or consider independently in the analysis need to be settled, based on the specific research questions to be addressed, and considering that rather complex relationships among “*exposure variables may exist and have the potential of, for example, introducing multi-collinearity problems during data analysis*” (38). An example is “age at starting” and smoking duration, two variables which are clearly not independent. Another example is the composite variable “pack years”; by integrating smoking duration and intensity, identical pack years can result from a large number of rather different smoking histories, and some of these histories have shown unusual patterns of dose-response (39). Also, some aspects of smoking history that are potentially relevant *per se* for the exposure-response relationship under investigation are not covered at all by “pack years”, including potential effects of timing and duration of periods of smoking cessation (38), mixed smoking or “polytobacco use” (i.e., smoking of different cigarettes or other products, like pipes or cigars (40)), as well as non-daily (i.e., occasional) smoking (20). Finally, the derived omnibus quantity “pack years” does not seem to comprehensively capture all of the impact of its two constituents, in particular intensity (e.g., (41)).

While the validity and reliability of the SQ has not yet been formally determined, it is aimed at capturing manifest behaviors rather than latent variables (e.g., intentions to quit or risk perceptions), which are more problematic to measure. Also, it is based on similar or even identical questions and response categories than those previously used in various other questionnaires in varying combinations and contexts, and it has been the guiding principle to adhere to general standards of designing exposure questionnaires for epidemiological studies (e.g., (42)). By applying the SQ categories to existing data sets, indicators of its measurement properties (including sensitivity, specificity, and positive predictive value) and its agreement with other instruments and definitions of smoking behavior can henceforth be determined, hopefully by a variety of independent researchers.

Unfortunately, like any measurement based on self-reports, the SQ suffers potentially from principal problems of misclassification and recall bias. Question 8 provides the most sensitive question of the SQ with regard to differential and non-differential recall bias. By applying the SQ repeatedly, in different populations (e.g., with and without smoking-related diseases), or through different modes of administration (e.g., self-administered, interviews), as well as by comparing the SQ measurements with those obtained through other instruments or with available information on past smoking behavior (e.g., pre-existing records), patterns and sources of over- and underreporting can be investigated. To enhance recall, methods providing an autobiographical framework can be deployed but reduce the standardization of the data-acquisition across respondents and can obviously be quite time-demanding. Also, they are likely to require computer-assistance for organizing complex information and managing visual prompts. While this is a rather different approach than what the SQ is aiming at, aided recall techniques (e.g., (43)), including providing photographs of different brands of cigarette packs (44), seem to be promising for collecting more

accurate retrospective information on past smoking behavior, when a quantification that is as detailed and accurate as possible is required.

This is, however, neither feasible nor intended with a questionnaire that focuses on brevity and practicality, and is probably not required for a wide range of research applications. In addition to the above pragmatic support for questionnaire-based data acquisition from a lack of practical alternatives, underreporting of tobacco use has been found to be “*relatively rare*” ((45), p. 175) and self-reported tobacco use has been assessed as “*accurate in most studies*” ((46), p. 1086) and “*highly consistent*” with urinary cotinine test results ((47), p. 1828). In a review of 29 studies based on designs and samples not suggesting systematic underreporting of smoking (as smoking cessation studies might, under certain conditions), LEE and FOREY (48) found half of the proportions of self-reported non-smokers whose cotinine levels indicated current smoking in the range of 1.9 to 4.5 percent, the median at 3.1 percent. They further reported that, based on studies with follow-up interviews, on average about 5 percent of self-reported never-smokers had reported ever smoking in a previous interview. Although the utility of collecting self-reported information on smoking has also more recently been supported (32), problems of imperfect reliability of smoking exposure measurements of the magnitude suggested by the above findings may be much more of a problem when the effects of environmental and other exposures on smoking-related diseases are studied, rather than those of active smoking; where these effects are relatively weak compared to those of cumulative active smoking, residual confounding can be a serious concern.

It should be appreciated that questionnaires have limitations, and that not all aspects of smoking-related exposure can be dealt with by collecting data based on the reports of smokers. Total smoking exposure is, in addition to the aspects described so far, also determined by smoking “topography” (49), namely the specific aspects of puffing, such as number of puffs per cigarette, puff volume, puff duration, and butt length (50), i.e., behaviors of which smokers are usually not fully and reliably aware. Another important exposure factor is inhalation habits: While pipe and/or cigar smokers appear to be able to accurately report whether or not they inhale their tobacco smoke (51), this seems not to be the case for cigarette smokers (52). Consequently, when current cigarette smoking topographies and inhalation habits are to be addressed, actual measurements are required, and there is in fact currently no scientifically rigorous way of measuring historic smoking topography and inhalation habits, as changes over time cannot be excluded, even when the type of cigarette smoked has not changed.

Even though some large population health surveys (including the US National Health and Nutrition Examination Surveys (NHANES) and the Canadian Health Measures Survey (CHMS)) measure smoking status through cotinine levels, very many studies are likely to rely on self-reports for the foreseeable future. Besides the effort and cost, biochemical measurements have other downsides, including their inability to differentiate between different sources of nicotine (product type), possible interferences through nicotine from non-combustible products, entanglement of



dose and time lag since exposure, as well as their transient nature and restriction to current (previous few hours for nicotine and 2–3 days for cotinine) smoking exposure when based on blood or urine sampling. Evaluating samples from hair or toenails, while in theory providing information about cumulative smoking exposure over the last few months, appears to require further development and currently available results show rather limited correspondence with self-reported exposure (53, 54). Overall, these limitations render biochemical measurements of very limited value for assessing the smoking history of individuals.

## COMPETING INTERESTS

RW, ZSW, and RD work for Philip Morris International (PMI), R&D, and hold shares of the company. GK previously worked for Philip Morris International (PMI) and holds shares of the company. He currently consults in the fields of statistics and epidemiology for a number of companies and academic institutions. CC, founder of Carson Watts Consulting, is an independent consultant in toxicology to a number of tobacco, pharmaceutical and chemical companies.

## AUTHORS' CONTRIBUTIONS

RW initiated the work and finalized the manuscript. GK and RW conceived the smoking questionnaire and drafted the manuscript. CC conducted comprehensive literature searches and wrote initial comprehensive versions of the manuscript. RD critically revised the manuscript at various stages and made substantial contributions to the contents and the text. ZSW conducted detailed literature analyses and reviewed and contributed to the manuscript and tables. All authors read and approved the final manuscript.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the comments made by Peter N. Lee and Barbara Forey on earlier versions of the manuscript.

The research described in this article was supported by Philip Morris International.

## REFERENCES

1. Doll, R. and A.B. Hill: Smoking and Carcinoma of the Lung. Preliminary Report; *Brit. Med. J.* 2 (1950) 739–748.
2. Wynder, E.L. and E.A. Graham: Tobacco Smoking as a Possible Etiologic Factor in Bronchiogenic Carcinoma. A Study of Six Hundred and Eighty-Four Proved Cases; *JAMA* 143 (1950) 329–336.
3. Massey, M.M., G. Boyd, M. Mattson, and M.R. Feinleib: Inventory of Surveys on Smoking; *Public Health Rep.* 102 (1987) 430–438.
4. US Department of Health and Human Services, Centers for Disease Control and Prevention: National Health Interview Survey Brochure and Questionnaires; available at: <http://www.cdc.gov/nchs/nhis.htm> (accessed August 01, 2013).
5. Bondy, S.J., J.C. Victor, and L.M. Diemert: Origin and Use of the 100 Cigarette Criterion in Tobacco Surveys; *Tob. Control* 18 (2009) 317–323.
6. Giovino, G.A.: Epidemiology of Tobacco Use in the United States; *Oncogene* 21 (2002) 7326–7340.
7. European Commission, Eurostat: European Health Interview Survey: EHIS Questionnaire; available at: [http://ec.europa.eu/health/ph\\_information/implementation/systems/docs/ev\\_20070315\\_ehis\\_en.pdf](http://ec.europa.eu/health/ph_information/implementation/systems/docs/ev_20070315_ehis_en.pdf) (accessed August 01, 2013).
8. World Health Organization (WHO): Guidelines for Controlling and Monitoring the Tobacco Epidemic. Part II. Monitoring the Tobacco Epidemic. Chapter 8. Monitoring Tobacco Use; WHO, Geneva, Switzerland, 1998.
9. International Agency for Research on Cancer (IARC): Tobacco Control: Reversal of Risk after Quitting Smoking; IARC, Lyon, France, 2007.
10. Delnevo, C.D. and U.E. Bauer: Monitoring the Tobacco Use Epidemic III: The Host: Data Sources and Methodological Challenges; *Prev. Med.* 48 (1 Suppl.) (2009) S16–S23.
11. Bogdanovica, I., F. Godfrey, A. McNeill, and J. Britton: Smoking Prevalence in the European Union: A Comparison of National and Transnational Prevalence Survey Methods and Results; *Tob. Control* 20 (2011) e4.
12. Okoli, C.T., P.A. Ratner, R.J. Haines, K.M. Sullivan, S.E. Guo, and J.L. Johnson: Do Researcher-Derived Classifications of Youths' Smoking Behavior Correspond with Youths' Characterizations of Their Behavior?; *Addict. Behav.* 34 (2009) 984–992.
13. Dietrich, T. and K. Hoffmann: A Comprehensive Index for the Modeling of Smoking History in Periodontal Research; *J. Dent. Res.* 83 (2004) 859–863.
14. Baker, R.R. and L.S. Lewis: A Review of the Incidence and Consequences of Cigarette Filter Vent Blocking Among Smokers; *Beitr. Tabakforsch. Int.* 19 (2001) 209–228.
15. Counts, M.E., F.S. Hsu, and F.J. Tewes: Development of a Commercial Cigarette "Market Map" Comparison Methodology for Evaluating New or Non-Conventional Cigarettes; *Regul. Toxicol. Pharmacol.* 46 (2006) 225–242.
16. Heck, J.D.: Smokers of Menthol and Nonmenthol Cigarettes Exhibit Similar Levels of Biomarkers of Smoke Exposure; *Cancer Epidemiol. Biomarkers Prev.* 18 (2009) 622–629.
17. Sarkar, M., S. Kapur, K. Frost-Pineda, S. Feng, J. Wang, Q. Liang, and H. Roethig: Evaluation of Biomarkers of Exposure to Selected Cigarette Smoke Constituents in Adult Smokers Switched to Carbon-Filtered Cigarettes in Short-Term and Long-Term Clinical Studies; *Nicotine Tob. Res.* 10 (2008) 1761–1772.
18. Gilpin, E., S.W. Cavin, and J.P. Pierce: Adult Smokers Who Do Not Smoke Daily; *Addiction* 92 (1997) 473–480.

19. Centers for Disease Control and Prevention (CDC): Prevalence of Current Cigarette Smoking Among Adults and Changes in Prevalence of Current and Some Day Smoking - United States, 1996–2001; *MMWR Morb. Mortal Wkly. Rep.* 52 (2003) 303–307.
20. Coggins, C.R., E.L. Murrelle, R.A. Carchman, and C. Heidbreder: Light and Intermittent Cigarette Smokers: A Review (1989–2009); *Psychopharmacology* 207 (2009) 343–363.
21. Doll, R. and R. Peto: Cigarette Smoking and Bronchial Carcinoma: Dose and Time Relationships Among Regular Smokers and Lifelong Non-Smokers; *J. Epidemiol. Community Health* 32 (1978) 303–313.
22. American Cancer Society: Cancer Prevention Study Overviews - Cancer Prevention Study II; available at [http://www.cancer.org/docroot/RES/content/RES\\_6\\_2\\_Study\\_Overviews.asp](http://www.cancer.org/docroot/RES/content/RES_6_2_Study_Overviews.asp) (accessed August 01, 2013).
23. Garfinkel, L.: Selection, Follow-up, and Analysis in the American Cancer Society Prospective Studies; *Natl. Cancer Inst. Monogr.* 67 (1985) 49–52.
24. Stellman, S.D. and L. Garfinkel: Smoking Habits and Tar Levels in a New American Cancer Society Prospective Study of 1.2 Million Men and Women; *J. Natl. Cancer Inst.* 76 (1986) 1057–1063.
25. Thun, M.J. and C.W. Heath Jr: Changes in Mortality from Smoking in Two American Cancer Society Prospective Studies Since 1959; *Prev. Med.* 26 (1997) 422–426.
26. Diemert, L.M., S.J. Bondy, J.C. Victor, J.E. Cohen, K.S. Brown, R. Ferrence, J. Garcia, P. McDonald, P. Selby, and T. Stephens: Efficient Screening of Current Smoking Status in Recruitment of Smokers for Population-Based Research; *Nicotine Tob. Res.* 10 (2008) 1663–1667.
27. Boffetta, P., F. Nyberg, A. Agudo, E. Benhamou, K.H. Jockel, M. Kreuzer, F. Merletti, G. Pershagen, H. Pohlbeln, L. Simonato, H.E. Wichmann, and R. Saracci: Risk of Lung Cancer From Exposure to Environmental Tobacco Smoke from Cigars, Cigarillos, and Pipes; *Int. J. Cancer* 83 (1999) 805–806.
28. Wood, D.M., M.G. Mould, S.B. Ong, and E.H. Baker: “Pack Year” Smoking Histories: What About Patients Who Use Loose Tobacco?; *Tob. Control* 14 (2005) 141–142.
29. Leffondré, K., M. Abrahamowicz, Y. Xiao, and J. Siemiatycki: Modelling Smoking History Using a Comprehensive Smoking Index: Applications to Lung Cancer; *Stat. Med.* 25 (2006) 4132–4146.
30. Neuner, B., J. Wellmann, B. Dasch, T. Behrens, B. Claes, M. Dietzel, D. Pauleikhoff, and H.W. Hense: Modeling Smoking History: A Comparison of Different Approaches in the MARS Study on Age-Related Maculopathy; *Ann. Epidemiol.* 17 (2007) 615–621.
31. Olsen, J. and IEA European Questionnaire Group: Epidemiology Deserves Better Questionnaires; *Int. J. Epidemiol.* 27 (1998) 935.
32. Brigham, J., C.N. Lessov-Schlaggar, H.S. Javitz, R.E. Krasnow, E. Tildesley, J. Andrews, H. Hops, M.D. Cornelius, N.L. Day, M. McElroy, and G.E. Swan: Validity of Recall of Tobacco Use in Two Prospective Cohorts; *Am. J. Epidemiol.* 172 (2010) 828–835.
33. Global Adult Tobacco Survey Collaborative Group: Tobacco Questions for Surveys: A Subset of Key Questions from the Global Adult Tobacco Survey (GATS), 2<sup>nd</sup> edition; Centers for Disease Control and Prevention (CDC), Atlanta, GA, USA, 2011.
34. Colby, S.M., M.A. Clark, M.L. Rogers, S. Ramsey, A.L. Graham, J. Boergers, C.W. Kahler, G.D. Papan-donatos, S.L. Buka, R.S. Niaura, and D.B. Abrams: Development and Reliability of the Lifetime Interview on Smoking Trajectories; *Nicotine Tob. Res.* 14 (2012) 290–298.
35. Fagerström, K.-O.: Measuring Degree of Physical Dependency to Tobacco Smoking with Reference to Individualization of Treatment; *Addict. Behav.* 3 (1978) 235–241.
36. Berg, C.J., E. Nehl, K. Sterling, T. Buchanan, S. Narula, E. Sutfin, and J.S. Ahluwalia: The Development and Validation of a Scale Assessing Individual Schemas Used in Classifying a Smoker: Implications for Research and Practice; *Nicotine Tob. Res.* 13 (2011) 1257–1265.
37. Ogden, M.W., W.T. Morgan, D.L. Heavner, R.A. Davis, and T.J. Steichen: National Incidence of Smoking and Misclassification Among the US Married Female Population; *J. Clin. Epidemiol.* 50 (1996) 253–263.
38. Leffondré, K., M. Abrahamowicz, J. Siemiatycki, and B. Rachet: Modeling Smoking History: A Comparison of Different Approaches; *Am. J. Epidemiol.* 156 (2002) 813–823.
39. Lubin, J.H., J. Virtamo, S.J. Weinstein, and D. Albanes: Cigarette Smoking and Cancer: Intensity Patterns in the alpha-Tocopherol, beta-Carotene Cancer Prevention Study in Finnish Men; *Am. J. Epidemiol.* 167 (2008) 970–975.
40. Bombard, J.M., L.L. Pederson, J.J. Koval, and M. O’Hegarty: How Are Lifetime Polytobacco Users Different Than Current Cigarette-Only Users? Results From a Canadian Young Adult Population; *Addict. Behav.* 34 (2009) 1069–1072.
41. Lubin, J.H. and N.E. Caporaso: Cigarette Smoking and Lung Cancer: Modeling Total Exposure and Intensity; *Cancer Epidemiol. Biomarkers Prev.* 15 (2006) 517–523.
42. Nieuwenhuijsen, M.J.: Design of Exposure Questionnaires for Epidemiological Studies; *Occup. Environ. Med.* 62 (2005) 272–280.
43. Reimer, M. and B. Matthes: Collecting Event Histories with Truetales: Techniques to Improve Autobiographical Recall Problems in Standardized Interviews; *Qual. Quant.* 41 (2007) 711–735.
44. Dijkstra, W., J.H. Smit, and Y.P. Ongena: An Evaluation Study of the Event History Calendar; *in: Calendar and Time Diary: Methods in Life Course Research*, edited by R.F. Belli, F.P. Stafford, and D.F. Alwin, Sage, Los Angeles, CA, USA, 2009, pp. 257–276.
45. Fendrich, M., M.E. Mackesy-Ami, T.P. Johnson, A. Hubbell, and J.S. Wislar: Tobacco-Reporting Validity in an Epidemiologic Drug-Use Survey; *Addict. Behav.* 30 (2005) 175–181.
46. Patrick, D.L., A. Cheadle, D.C. Thompson, P. Diehr, T. Koepsell, and S. Kinne: The Validity of Self-Reported Smoking: A Review and Meta-Analysis; *Am. J. Public*

- Health 84 (1994) 1086–1093.
47. Studts, J.L., S.R. Ghate, J.L. Gill, C.R. Studts, C.N. Barnes, A.S. LaJoie, M.A. Andrykowski, and R.V. LaRocca: Validity of Self-reported Smoking Status Among Participants in a Lung Cancer Screening Trial; *Cancer Epidemiol. Biomarkers Prev.* 15 (2006) 1825–1828.
  48. Lee, P.N. and B. Forey: Misclassification of Smoking Habits as Determined by Cotinine or by Repeated Self-Report - A Summary of Evidence from 42 Studies; *J. Smoking Rel. Dis.* 6 (1995) 109–129.
  49. Hammond, D., G.T. Fong, K.M. Cummings, and A. Hyland: Smoking Topography, Brand Switching, and Nicotine Delivery: Results From an *in Vivo* Study; *Cancer Epidemiol. Biomarkers Prev.* 14 (2005) 1370–1375.
  50. St. Charles, F.K., A.A. Kabbani, and M.W. Ogden: Smoking Behavior: How Close to the Tipping Do Consumers Actually Smoke?; *Beitr. Tabakforsch. Int.* 21 (2005) 435–440.
  51. Herling, S. and L.T. Kozlowski: The Importance of Direct Questions About Inhalation and Daily Intake in the Evaluation of Pipe and Cigar Smokers; *Prev. Med.* 17 (1988) 73–78.
  52. Stepney, R.: Are Smokers' Self-Reports of Inhalation a Useful Measure of Smoke Exposure?; *J. Epidemiol. Community Health* 36 (1982) 109–112.
  53. Stepanov, I., S.S. Hecht, B. Lindgren, P.J. 3<sup>rd</sup>, M. Wilson and N.L. Benowitz: Relationship of Human Toenail Nicotine, Cotinine, and 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanol to Levels of these Biomarkers in Plasma and Urine; *Cancer Epidemiol. Biomarkers Prev.* 16 (2007) 1382–1386.
  54. Al-Delaimy, W.K. and W.C. Willett: Measurement of Tobacco Smoke Exposure: Comparison of Toenail Nicotine Biomarkers and Self-Reports; *Cancer Epidemiol. Biomarkers Prev.* 17 (2008) 1255–1261.

*Corresponding author:*

*Rolf Weitkunat  
PMI Research & Development  
Philip Morris Products SA  
Neuchâtel, Switzerland  
E-mail: rolf.weitkunat@pmi.com*

*Other E-mail addresses:  
chris@carsonwattsconsulting.com  
zheng.wang@pmi.com  
ruth.dempsey@pmi.com  
gerd.kallischnigg@gmx.de*