Title
Cotinine validation of self-reported smoking in commercially run community surveys

Journal
Journal of Clinical Epidemiology, 40(7)

ISSN
0895-4356

Authors
Pierce, JP
Dwyer, T
DiGiusto, E
et al.

Publication Date
1987

DOI
10.1016/0021-9681(87)90105-6

Peer reviewed

J Chron Dis Vol. 40, No. 7, pp. 689–695, 1987
Printed in Great Britain.

0021-9681/87 $3.00 + 0.00

COTININE VALIDATION OF SELF-REPORTED SMOKING IN COMMERCIALLY RUN COMMUNITY SURVEYS

J. P. PIERCE, T. DWYER, E. DIGIUSTO, T. CARPENTER, C. HANNAM, A. AMIN, C. YONG, G. SARFATY, J. SHAW, N. BURKE and QUIT FOR LIFE STEERING COMMITTEE*

School of Public Health and Tropical Medicine, University of Sydney, NSW 2006, Australia

(*Received in revised form* 11 *August* 1986)

Abstract—A validation study was carried out on self-reported smoking for 1177 people in Sydney and Melbourne in 1983. Because of its long half life and the fact that smoking is its only source in body fluids, saliva cotinine was chosen as the validation measure. Cotinine levels above 250 nmol/l were used to classify people as smokers. The sensitivity of self-reported smoking was 92.6% and the specificity was 93.4%. There was some evidence that people in the process of changing their smoking status might be slow in updating their self-classification. The smoking prevalence estimate based on cotinine levels was found to be 1.7% lower than that for self-reported smoking status. The small proportion of false negatives and false positives suggests that commercially collected data banks can be valid sources of prevalence data. Correlation between cotinine level and reported cigarette consumption was not affected by sample volume, and was similar to that achieved for carbon monoxide and thiocyanate at a low 0.34. Regression analysis using self-reported cigarette consumption filter/non-filter cigarettes, and time since last cigarette as predictors, explained 13.6% of the variance in cotinine level.

Smoking Validation Cotinine Self-report Surveys Epidemiology

INTRODUCTION

Many public health researchers have strongly suggested that biochemical markers of cigarette smoking behaviour be used to validate self-reported smoking [1–4]. The argument has been particularly directed at evaluations of both cessation and prevention programs where the increasing social unacceptability of smoking could result in significant under-reporting [5–7]. However, studies of patients in several settings including a pre-paid medical practice, a large randomized trial and a large community study [2, 8, 9] have demonstrated the accuracy of self-reported smoking status. The validity of self-reported smoking status is essential to the usefulness of smoking data collected in historical data banks if they are to be used to analyze trends in community behaviour. This study set out to assess the validity of one major set of routinely collected data on self-report smoking status.

In Australia, a large commercial research company has been regularly collecting community-wide smoking data since the mid-1960's These data include demographic variables and their availability provides opportunity to plot trends over time within sub-categories of respondents. Such information is useful to public health workers seeking to target their efforts to reduce the prevalence of cigarette smoking.

*The project was managed by a Steering Committee: A. Cripps (Chairperson),[1] J. Carson,[1] T. Dwyer,[2] G. Frape,[1] D. Gadiel,[3] E. Henry,[4] B. Herriot,[5] B. Higham,[1] J. Mullins,[6] J. Pierce,[2] C. Sarfaty,[4] J. Shaw[7] and S. Walker.[7]
[1]New South Wales Department of Health, [2]Commonwealth Institute of Health, [3]Hospitals Contribution Fund of Australia, [4]New South Wales State Cancer Council, [5]Australian Medical Association, [6]Pharmacy Guild of New South Wales and [7]National Heart Foundation of Australia.

Such data also provide opportunities for regular comparable evaluation of the effectiveness of various public health efforts in different communities. These data are being used to evaluate the effectiveness of the "Quit for Life" anti-smoking campaign which took place in Sydney during the winter of 1983.

Validation of self-reported smoking can be done by examining concentrations of carbon monoxide, thiocyanate or cotinine in either body fluids or expired air. Cigarette smoking is not the only source of either carbon monoxide or thiocyanate in body fluids [10, 11], and hence examination of their concentration levels is not the ideal validatory measure. Cotinine however, is a direct metabolite of nicotine, and tobacco is the only known source of the substance in our population (Nicorette chewing gum had not been released at the time of our study). Cotinine has a reported half-life of 20–25 hours [12, 13] and is stable across and within body fluids with a recorded 10% variation over the day [13–15]. These characteristics reduce the need for controlling the time of day of sample collection and the time between the last cigarette and sample collection. Extremely low concentrations of cotinine can be detected (less than 100 nmol/l), however, at these extreme levels, laboratory quality control can be a problem. A recent study [9] with cotinine as the biochemical validator of self-reported smoking used a decision point of 250 nmol/l to classify people as smokers or non-smokers. Using self reported smoking status as the standard, they reported a sensitivity of plasma continine of 98% and a specificity of 95%. Plasma has been the preferred fluid for continine determination, however, a good correlation has been reported between plasma and saliva cotinine levels and saliva cotinine has been shown to be an effective tool for measurement of smoking prevalence [20, 21].

METHODS

Smoking questions are included in the weekly national survey procedure of the Roy Morgan Research Company, one of the largest such companies in Australia. This procedure involves the use of trained interviewers who follow a standard protocol for conducting door knock surveys from 9 a.m. to 4 p.m. on Saturdays and Sundays on most weekends of the year. The organization has built-in quality control procedures and the nationwide response rate to all houses visited was 60%. More details of the response rate are contained in an associated paper [16].

A two-tiered random sampling procedure is used, starting with the selection of electoral sub-divisions, and followed by the selection of an address from within that sub-division. Interviewers begin at this address and knock on consecutive houses in a clockwise direction until they obtain 10 interviews. There is a single call-back made at each non-respondent address. The standard procedure involves randomly choosing a sex for the first address and requesting an interview with the youngest person over 14 of that sex in the house. The choice of sex is then alternated in ensuing interviews until a maximum of five interviews of the cluster of 10 are obtained for any one sex.

The initial section of the interview consisted of a 20 minute questionnaire, which included several questions about the respondents smoking habits. All respondents were handed a card with ten categories of smoking status, and were asked to choose the statement that best described themselves. The first six choices were different combinations of smoking cigarettes, cigars and pipes. The next three related to being an ex-smoker of cigarettes, cigars or pipes. The remaining classification was never smoked at all. In addition, smokers were asked how many cigarettes they smoked each day (see Appendix). On two consecutive weekends in May 1983, following completion of the normal questionnaire, 1172 people in Sydney and Melbourne were asked to provide a saliva sample for a Department of Health analysis related to environmental pollution. It should be noted that at the time of completing the questionnaire, respondents were unaware that any further tests of their smoking status would take place. As far as possible, interviewers tried to ensure that samples were collected at least 5 minutes after any eating, drinking or smoking.

Each respondent was handed an air-tight plastic tube containing a dental cotton roll. S/he was asked to place the cotton roll in the mouth and chew gently for a minute until the roll was saturated with saliva. The roll was deposited directly from the mouth into the tube without contact with the fingers. After collection, samples were stored on ice in a polystyrene container for a few hours and were subsequently stored frozen until assayed. At the time of assay, as much saliva as possible was squeezed from each cotton roll using disposable syringes placed

Table 1. Description of inclusions and non-inclusions in the validation study

		Non-Inclusions		Inclusions	
		%	(*N*)	%	(*N*)
Sex	Male	45.2	(89)	50.8	(495)
	Female	54.8	(108)	49.2	(480)
Age	14–19	13.7	(27)	14.9	(145)
	20–39	40.6	(80)	46.5	(453)
	40–64	26.4	(52)	28.8	(281)
	65+	19.3	(38)	9.8	(96)
Education	Less than HSC*	68.5	(135)	64.7	(631)
	HSC or equivalent	10.7	(125)	10.7	(32)
	Tertiary	15.2	(30)	15.7	(153)
Smoking status	Smoker	33.0	(65)	36.2	(353)
	Ex-smoker	19.8	(39)	21.2	(207)
	Non-smoker	46.7	(92)	41.9	(409)
Locality	Sydney	67.5	(133)	54.5	(531)
	Melbourne	32.5	(64)	45.5	(444)
Total		16.8	(197)	83.2	(975)

*HSC is the High School Certificate and the university entrance qualification in Australia.

in a mechanical press. Saliva levels of cotinine were determined using 12.5 × 0.2 mm carbowax capillary mounted in a Hewlett-Packard 5880A gas chromatograph.

RESULTS

Non-respondents

Conclusions from studies are often biased because different response rates are obtained from the various sub-groups of the study population. In this study, data are available on those people who responded to the overall survey, but who declined to give a saliva sample or provided a sample with a volume too small for analysis. Table 1 compares those included in the analysis with those not included, either because they were non-respondents (150 cases) or because their samples were lost in the analysis procedure (47 cases). The non-inclusion rate for the smoking validation sub-study was 16.8% and there were differences in the characteristics of people included in the study compared to those not included. A higher proportion of women, people over 65 years of age, self-reported non-smokers and residents of Sydney were in the non-inclusion group. The difference in response rate between the two cities is only slightly higher than would be expected from the known proportion of people at home during a weekend and the regular difference in refusal rate between the two cities [17]. The greater proportion of self-reported non-smokers among the non-respondents represents a potential bias to the study findings.

The cotinine decision level

The precision of the laboratory procedure was checked by reproducing the results of Pojer *et al.* [9] on the number of smokers misclassified at varying decision levels of cotinine (see Table 2).

Misclassifications refer to self-reports of smoking when the cotinine level indicated non-smoking and self-reports of non-smoking when the cotinine level indicated smoking. The data broadly support the earlier findings, and 250 nmol/l was used as the decision point for classifying a person as a smoker.

As 40% of our sample had a retrievable saliva volume under 0.6 ml, it was necessary to know whether there was a minimum saliva volume

Table 2. Classifications of self-reported smoking for different levels of cotinine

Cotinine decision level* (nmol/l)	Percent correctly classified		Percent total misclassifications of self-report
	Smokers	Non-smokers	
0	93.5	53.1	53.4
50	92.1	82.6	25.4
100	90.7	90.5	18.8
150	90.1	93.6	16.3
200	89.0	95.0	16.0
250	88.1	96.0	15.9
300	87.8	96.6	15.6
350	85.6	96.8	17.6
400	81.3	97.6	21.1
sample size	(353)	(622)	

*The level of cotinine above which subjects are classified as smokers. Total misclassifications refers to the combined proportion of self-reported smokers when saliva cotinine indicated non-smoking and self-reported non-smokers when saliva cotinine indicated smoking.

Table 3. Self-reported vs cotinine-assessed smoking status

		Cotinine-assessed:		
		Smoker	Non-smoker	Total
Self-reported:	Smoker	311	42	353
	Non-smoker	25	597	622
	Total	336	639	975

that should be used in the study. The minimum volume decided upon was 0.05 ml, since there was no alteration in the number of misclassifications in volumes greater than this level.

Sensitivity and specificity

Table 3 presents the data differences between cotinine (used here as the gold standard) and self-reported classification of smoking status. There were 25 people who had a smoking cotinine level, who described themselves as non-smokers (92.6% sensitivity), and 42 people with non-smoking cotinine levels who reported themselves as smokers (93.4% specificity).

Of the 42 self-reported smokers assessed by cotinine to be non-smokers, five indicated that they only smoked a cigar or pipe and a further five reported not having smoked for at least 24 hours. Both of these reasons are consistent with a low cotinine reading, and thus should not be considered as errors. If the remaining 32 are considered the only real false positives the specificity increases to 95.0%. Half of the remainder indicated that they were extremely likely to quit smoking in the short term compared to only 8% of the general smoking population. This could indicate that some people still labelled themselves as smokers although they were in the process of quitting.

Of the 25 false negatives, that is those self-reported non-smokers assessed by cotinine to be smokers, 10 stated they had never smoked, and 15 indicated that they were ex-smokers. Of these self-reported ex-smokers, four indicated that they had recently quit, and three of these felt that a relapse was likely in the short term. It is possible that this group had recently relapsed, but were as yet unwilling to internally re-classify themselves as smokers.

Cotinine level and reported cigarette consumption

The correlation between cotinine levels and reported cigarette consumption for all cigarette smokers was 0.325. This improved to 0.344 if the sample was restricted to only those with a saliva volume over 0.4 ml.

This correlation should be affected by the amount of nicotine in the cigarette that is smoked, the degree of inhalation in the smoking technique, the time since the last cigarette, as well as the biochemical uptake and metabolism of the individual. This study only had measures of two of these variables: time since the last cigarette, and a crude measure of cigarette strength (filter/non-filter). A regression analysis was undertaken including these variables with reported cigarette consumption as predictors of the cotinine level, and the model that resulted is presented in Table 4. As expected, the strongest predictor was the number of cigarettes smoked per day, but all three variables were important in the model. The amount of variation explained by this model was a low 13.6%. The lack of explanatory power of this model could be related to the lack of information on smoking style or metabolism, the variability of the cotinine measure itself, and to using only a dichotomous variable for cigarette strength.

Cotinine validation compared to thiocyanate and carbon monoxide

Fortmann *et al.* [3] have published descriptive data on self-reported smoking, expired air carbon monoxide and plasma thiocyanate levels for participants in the baseline surveys of the Stanford Five Cities Project. These data are reproduced along with similar data from our study in Table 5. The proportion of the population in each self-reported smoking category and the self-reported cigarette consumption per day within the categories are remarkably similar for people from large Californian towns and Australian capital cities. The similarity between the mean number of cigarettes smoked, and the

Table 4. Regression model of predictor variables for cotinine levels in smokers

Predictor	Coefficient	SE	t	p
Cigarettes/day	25.97	5.50	4.7	<0.0001
Time since last cig.	−0.72	0.28	−2.6	<0.02
Cigarette strength	140.08	64.04	2.2	<0.03

($N = 353$, $R^2 = 13.6$).

Table 5. Comparisons of study with Stanford data* on smoking measures

	Self-reported cigarettes/day		Validation		
	Quit for Life	Stanford Project	Cotinine†	CO‡	SCN§
Regular smoker	(>9 cigs/day)				
Percent of sample	27.3	25.4			
Mean cigs/day	22.4	22.5	1964.9	27.3	13.7
Light smoker	(<9 cigs/day)				
Percent of sample	6.1	5.7			
Mean cigs/day	4.3	3.3	1016.0	11.1	89.4
Ex-smoker					
Percent of sample	22.3	22.0			
			97.7	5.0	56.5
Non-smoker					
Percent of sample	44.0	45.4			
			41.7	4.6	53.1

*Fortmann *et al.* [3].
†Cotinine in nmol/l ("Quit for Life").
‡Carbon monoxide in ppm (Stanford project).
§Thiocyanate in mmol/l (Stanford project).

percent of the population in the different smoking categories, allows comparisons to be made between these different validation measures on their ability to discriminate people who only smoke a small amount.

Using logistic regression, Fortmann *et al.* [3] graphed the probability of being labelled a smoker by either the thiocyanate or the carbon monoxide validations, by the self-reported number of cigarettes smoked per day. They noted that most self-reported light and irregular smokers were misclassified by one or both of these procedures. Using the same satistical procedure, we compared the results of the cotinine classifications by number of cigarettes smoked per day to the results obtained in the Stanford study (Fig 1). Comparison of the logistic curves indicates that cotinine could be a superior discriminator at very low levels of smoking; while carbon monoxide seems to have a very slight advantage once people report smoking ten or more cigarettes per day.

DISCUSSION

The results from this study indicate that if saliva cotinine was used as the true measure of smoking status instead of self-report, there would be a downward adjustment in smoking prevalence of 1.7%. This small discrepancy between the two measures—neither of which provides incontrovertible evidence of smoking status—did not, it was decided, necessitate

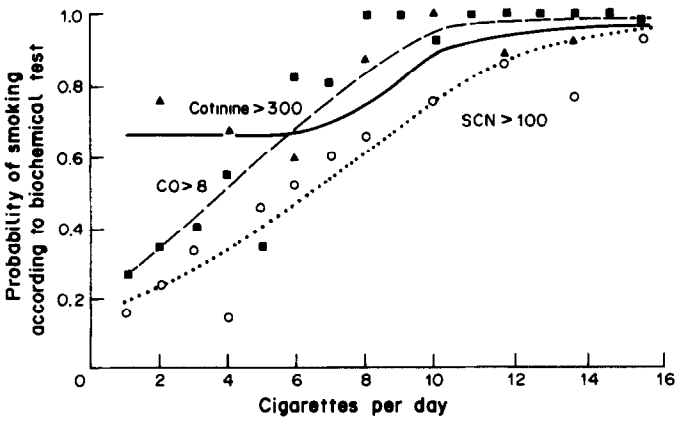

Fig. 1. Fitted logistic probability of being classified as a smoker according to cotinine (——), CO (– – –) and SCN (·····) by level of self-reported smoking. Also shown are the actual proportions of the people reporting each smoking level who would be classified as smokers by cotinine (▲), CO (■) and SCN (○). CO and SCN data from Fortmann *et al.* [3].

adjustment of the self-reported smoking data collected from community samples.

The total non-inclusion rate of 17% was evenly distributed across major demographic and smoking categories, although the over 65 year age group were under-represented as respondents. The known decrease in saliva volume with age could be one reason for this. As this age group is rarely the major interest group in smoking research, their under-representation in this study is not expected to seriously bias the results. Saliva volumes in this study tended to be lower than thought optimal, but this had little effect on the cotinine classification of smokers and non-smokers.

This study demonstrated that the sensitivity and specificity for self-reported smoking were 92.6% and 93.4% respectively. Analysis of the smoking behaviour of those with a non-smoking cotinine reading, who reported themselves as smokers, indicates that the false positive rate was most probably artificially high because of the inclusion of pipe/cigar smokers and people who had not smoked in the last 24 hours. Removal of these two groups increased specificity to 95.0%

Analysis of false positives (i.e. those classified by the cotinine level as non-smokers who reported themselves as smokers) suggested a possible quit-group of people whose self-perception of their smoking status lagged behind their success in giving up smoking. Similarly, analysis of false negatives (cotinine classified smokers but self-reported non-smokers) indicated that certain, recent recidivists had not yet altered their self-perception back to "smoker". Accuracy in self-reported smoking status could probably be increased if the all-or-none nature of smoking status categories was changed to allow for an intermediate category, for those in the process of either giving up or starting to smoke.

Saliva cotinine as a validation procedure for self-reported smoking was compared with results on thiocyanate and carbon monoxide from a previous study [3]. The proportions in each self-reported smoking category for both study populations were remarkably similar, given that the studies were done in different sub groups in different countries. The range for saliva cotinine is much greater than the other two measures and it appears to be a better predictor of smoking status with those who smoke only a few cigarattes a day.

The correlation between cotinine level and reported cigarette consumption is of the order reported for both carbon monoxide and thiocyanate validation analyses, which must be regarded as less than optimal. However, the correlation between cotinine level and cigarette consumption was comparable to previous reports. Benowitz *et al.* reported a correlation of 0.4 and Hill *et al.* reported a correlation of 0.45 among men and 0.39 among women [18, 19]. A regression analysis which took account of the time since the last cigarette was smoked, whether or not the cigarette had a filter, and reported cigarette consumption, explained 13.6% of the variation in cotinine levels. This low result could be due to inaccurate reporting of the number of cigarettes smoked, the variability in strength of cigarettes, the cotinine measure itself and smoking style or differences in metabolism. In the "Quit for Life" follow-up study, more detailed information will be available on the nicotine strength of cigarettes smoked, and it is hoped that this will increase the amount of variation explained by this model. Saliva cotinine, according to comparison with the Stanford data, appears to be a better validation measure of self-reported smoking at low levels of cigarette consumption than either carbon monoxide or thiocynate. However, the introduction onto the market of "nicorette" chewing gum since our study represents a potential confounder, as cotinine levels could conceivably be high for someone who has not had a cigarette, but who has been chewing gum. Obviously this would lead to a higher percentage of false negatives being recorded. Given also the comparatively high cost of biochemical analysis (approximately $15 per sample for cotinine), it would seem that carbon monoxide, where an electrochemical sensor can be used repeatedly on multiple samples of expired air, is the preferred validation measure in community surveys.

A general cause for concern in self-report community surveys has been the hypothesized tendency of smokers to report non-smoking status. This study found no evidence of a significant misclassification in this direction. Indeed, the objective measure of smoking status gave a lower estimate of the proportion of the population who smoked, although the discrepancy (1.7%) was not so large as to warrant adjustments to the data. Accordingly, supplementary validation of smoking status is not indicated by this study to be crucial to large, cross-sectional community surveys. Further validation of self reports must still be considered

necessary, however, in longitudinal studies of change in smoking behaviour.

Acknowledgements—The Quit for Life campaign was conceived by the Secretary of the New South Wales Health Department. Additional support was received from the Australian Department of Health, the Hospitals Contribution Fund of Australia and the New South Wales State Cancer Council.

REFERENCES

1. Cohen JD, Bartsch GE: A comparison between carboxyhaemoglobin and serum thiocyanate determinations as indicators of cigarette smoking. **Am J Public Health** 70: 284–286, 1980
2. Evans RI, Hansen WB, Mittlemark MB: Increasing the validity of self-reports of smoking behaviour in children. **J Appl Psychol** 62: 521–523, 1977
3. Fortmann SP, Rogers T, Vranizan K, Haskell WL, Solomon DS, Farquhar JW: Indirect measures of cigarette use: expired-air carbon monoxide versus plasma thiocyanate. **Prev Med** 13: 127-134, 1984
4. Feyerband C, Russell MAH: Rapid gas-liquid chromatographic determination of cotinine in biological fluids. **Analyst** 105: 998–1001, 1980
5. Hill P, Marquardt H: Plasma and urine changes after smoking different brands of cigarettes. **Clin Pharmacol Ther** 27: 652–658, 1980
6. Langone JJ, Van Vunakis H, Hill P: Quantification of cotinine in sera of smokers. **Res Commun Chem Pathol Pharmacol** 10: 21–28, 1975
7. Luepker RV, Pechacek TF, Murray DM, Johnson CA, Hund F, Jacobs DR: Saliva thiocyanate: A chemical indicator of cigarette smoking in adolescents. **Am J Public Health** 71: 1320–1324, 1981
8. Petitti DB, Friedman GD, Kahn W: Accuracy of information on smoking habits provided on self-administered research questionnaires. **Am J Public Health** 71: 308–311, 1981
9. Pojer R, Whitfield JB, Poulos V, Eckhardt IF, Richmond R, Hensley WJ: Comparison of carboxyhaemoglobin, cotinine and thiocyanate as methods to distinguish between smokers and non-smokers. **Clin Chem** 30, 1984
10. Prue DM, Martin JE, Hume AS: A critical evaluation of thocyanate as a biochemical index of smoking exposure. **Behav Ther** 11: 368–379, 1980
11. Russell MAH, Jarvis M, Iyer R, Feyerband C: Relation of nicotine yield of cigarettes to blood nicotine concentrations in smokers. **Br Med J** 1: 972–976, 1980
12. Sillet RW, Wilson MB, Malcolm RE, Ball KP: Deception among smokers. **Br Med J** 2: 1185–1186, 1978
13. Vogt TL, Selvin S, Widdowson G, Hulley SB. Expired air carbon monoxide and serum thocyanate as objective measures of cigarette exposure. **Am J Public Health** 67: 545–549, 1977
14. Warner KE: Possible increases in the under-reporting of cigarette consumption. **J Am Stat Assoc** 73: 314–318, 1973
15. Wilcox RG, Hughes J, Roland J: Verification of smoking history in patients after infarction using urinary nicotine and cotinine measurements. **Br Med J** 2: 1026–1028, 1979
16. Dwyer T, Pierce JP, Hannam CD, Amin A: Evaluation of Sydney "Quit For Life" anti-smoking campaign: 2. Changes in smoking prevalence **Med J Aust** 144: 344–347, 1986
17. National Heart Foundation of Australia: Risk factor prevalence study. Report No 1. National Heart Foundation, Canberra 1980
18. Benowitz NL, Hall SM, Herning RI, Jacob P, Jones RT, Osman A: Smokers of low-yield cigarettes do not consume less nicotine. **N Engl J Med** 309: 139–142, 1983
19. Hill P, Haley NJ, Wynder EL: Cigarette smoking: carboxyhaemoglobin, plasma nicotine, cotinine and thiocyanate vs self-reported smoking data and cardiovascular disease. **J Chron Dis** 36: 439–449, 1983
20. Sepkovic DW, Haley NJ: Biomedical applications of cotinine quantitation in smoking related research. **Am J Public Health** 75: 663–665, 1985
21. Di Giusto E, Eckhard I. Some properties of saliva cotinine measurements in indicating exposure to tobacco smoking. **Am J Public Health** 76: 1245–1246, 1986

APPENDIX

Smoking Categorisation Scheme

1. Regularly smoke only cigarettes.
2. Regularly smoke only cigarettes and cigars/pipe.
3. Regularly smoke cigars, used to smoke cigarettes.
4. Regularly smoke pipe, used to smoke cigarettes.
5. Regularly smoke cigars, never smoked cigarettes.
6. Regularly smoke pipe, never smoked cigarettes.
7. Don't smoke now, used to smoke only cigarettes.
8. Don't smoke now, used to smoke cigarettes and cigars/pipe.
9. Don't smoke now, used to smoke only cigars/pipe.
10. Never smoked at all.