Web-based and Mailed Questionnaires: *A Comparison of Response Rates and Compliance*

Katarina Augustsson Bälter, Olle Bälter, Elinor Fondell, and Ylva Trolle Lagerros

Background: We assessed response rates and compliance for a printed questionnaire and a Web questionnaire in a Swedish population-based study and explored the influence of adding personalized feedback to the Web questionnaire.

Methods: We assigned 875 subjects to 1 of 3 groups: printed questionnaire, plain Web questionnaire, or Web questionnaire with personalized feedback. The questionnaire had 2 parts, first a general section and then a dietary section.

Results: The response rate for the general section was 64% for the printed questionnaire, compared with 50% for the Web questionnaire with feedback. For the dietary questionnaire, the rates were reversed, resulting in a total response rate for the dietary questionnaire that did not differ between printed and web questionnaire with feedback.

Conclusions: Interactivity in the Web questionnaire increased compliance in completion of the second section of the questionnaire. Web questionnaires can be useful for research purposes in settings in which Internet access is high.

Submitted 18 August 2004; final version accepted 1 March 2005.

From the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

Supported by the Swedish Cancer Society and the Swedish Council for Working Life and Social Research.

Supplemental material for this article is available with the online version of the journal at www.epidem.com

Correspondence: Katarina Augustsson Bälter, Department of Medical Epidemiology and Biostatistics, Nobels väg 12A, Karolinska Institutet, SE- 171 77 Stockholm, Sweden. E-mail: katarina.balter@meb.ki.se.

Copyright © 2005 by Lippincott Williams & Wilkins

DOI: 10.1097/01.ede.0000164553.16591.4b

The Internet is an unexploited resource for collecting self-reported information in epidemiologic studies. Web-based questionnaires are easy to administer and offer several advantages, including immediate checks for incomplete or implausible answers, reminder messages to the respondent, automatic summarization of answers, personalized feedback, inclusion of illustrations or sounds to clarify complex questions, and hiding nonrelevant follow-up questions. Web questionnaires require no expense for printing, postage, manual check of incomplete answers, and transfer of data to an electronic format. The major cost for Web questionnaires is development of the system for handling the questionnaires; thus, once the system is established, the extra cost to add a few thousand or even a few hundred thousand participants to the study is relatively small.

However, access to the Internet can be biased with regard to age, sex, and education, among other factors. This differential access has restricted the use of Web-based methods to studies in specific groups with access to Internet, such as university students,[1] employees at certain companies,[2] or known Internet users.[3,4] Also, substantial proportion of the general population is unfamiliar with Web questionnaires and might hesitate to answer a Web questionnaire due to lack of experience or worries about security issues. However, these obstacles are likely to diminish over time. We explored response rates and compliance in a population-based study in Sweden, in which Internet access is estimated to be 80% in the working population.

METHODS

Study Population and Design

The study base comprised all persons 20 to 59 years of age living in a middle-sized county in Sweden in 2002. The county has the same distribution of inhabitants living in city (80%) and rural areas (20%) as the average in Sweden. We randomly selected 875 eligible persons from the Swedish Population Registry. Participation required filling out a questionnaire and undergoing several 24-hour recall interviews about physical activity. There were 3 versions of the questionnaire: (1) traditional printed questionnaire, (2) regular Web questionnaire, and (3) interactive Web questionnaire with personalized feedback.

The eligible sample was assigned randomly at the outset to one of the 3 versions of the questionnaires. All were sent an invitation letter informing them about the study. After 2 weeks, one third of the group was sent a printed questionnaire and two thirds were sent a letter with information on how to access the Web questionnaire, including details on use of the Web browser, the URL to our Web questionnaire, and an individual username. In addition, half of the Web group (one third of the total) was given the option of personalized feedback about their energy expenditure and body mass index. All nonrespondents were reminded after 3 weeks by a letter and contacted by phone after an additional 3 weeks, if necessary.

The questionnaire was divided into 2 parts. The first part was a general survey of lifestyle factors, such as physical activity, weight, height, smoking, and education, as well as the respondent's Internet habits. On the last page of this section, respondents were asked if they would be willing to answer additional questions about diet. Web respondents answering "yes" were linked to the second part of the questionnaire, whereas respondents to the printed questionnaire were sent the dietary questionnaire by returning mail. Personalized feedback was given to the interactive group on their intake of fiber, calcium, vitamin C, iron, and the composition of lunch and dinner meals. All basic questionnaires were followed by a validation study on physical activity.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

The ethical committee at the Karolinska Institutet approved the study. Answering the questionnaire was considered to be informed consent.

Technical Aspects of the Web Questionnaire

We collaborated with a commercial Swedish Web survey company (Netsurvey[2]), which developed software to meet our requirements with respect to layout, feedback, and interactivity. An individual username for login provided the same identification in the Web questionnaires as in the printed questionnaires. This username prevented multiple answers from the same respondent or answers from individuals other than the invited respondents and allowed us to direct reminders to the nonrespondents only. The respondent used the username to establish an encrypted connection, using Secure Socket Layer, with the Web questionnaire system at Netsurvey. The username was linked to a unique number that identified the person and the study. The questionnaire answers were stored at Netsurvey along with the unique identification number.

To minimize typing errors by the respondent, the system immediately checked for implausible answers (eg. letters instead of numbers, unrealistic weight or height, or missing answers). When the system discovered an error, the respondent was given the chance to change the answer before moving on to the next page.

RESULTS

Among the subjects given the printed questionnaire, the response rate for the general section was 64% (Table 1). Response rates for the groups using Web-based questionnaires were 51% for the questionnaire without feedback and 50% with feedback. Additional details regarding response rates can be found in a supplementary table, available with the electronic version of this article.

Compliance (willingness to answer the second part of the questionnaire) was higher for the Web questionnaires than for the printed questionnaire. Fifty-three percent of those who completed the general section of the paper questionnaire went on to complete the dietary part, compared with 58% of those responding to the Web questionnaire and 64% of those who were given the interactive Web questionnaire. Thus, the total response rate for the dietary questionnaire was similar for the 3 groups (34%, 29 and 32% respectively).

There were no differences in response to the 3 questionnaire options by age, body mass index, and current smoking, and only small differences by sex, education, and food habits (Table 2). The self-reported time spent answering the questionnaires did not differ between the groups. In general, those responding to the Web questionnaires were also more frequent users of the Internet and more able to arrange for privacy when using the Internet, assuring that no one was watching the screen.

DISCUSSION

To the best of our knowledge, this is the first population-based study comparing the use of Web questionnaires with a similar printed questionnaire. Half of the people invited to use the Web questionnaire completed the general section of questionnaire. The willingness to answer a second part of the questionnaire was higher with the Web questionnaire than with the printed questionnaire, which suggests that those that responding to the Web questionnaire found the process more appealing than those who responded to the mailed questionnaire.

Web questionnaires can be used for research purposes in population-based settings in which Internet access is high, although we found that the initial response rate was lower than for the traditional printed questionnaire. In comparison, the willingness to answer a second questionnaire was higher

TABLE 1. Response Rates for Each Group by Type of Questionnaires; Printed Questionnaire, Basic Web-based Questionnaire, and Web Questionnaire With Personalized Feedback

	Printed Questionnaire (n = 292)	Web Questionnaire (n = 293)	Web Questionnaire With Feedback (n = 290)
Number of subjects answering the first section of the questionnaire*	188	149	146
Response rate; % (95% CI)	64 (59–70)	51 (45–57)	50 (45–56)
Number of subjects continuing with the second section of the questionnaire†	100	86	93
Continuation rate; % (95% CI)‡	53 (46–60)	58 (50–66)	64 (56–71)
Total response rate for the second section; % (95% CI)	34 (29–40)	29 (24–35)	32 (27–37)

*The general section of the questionnaire about lifestyle factors (the first part of the questionnaire).
†The food frequency questionnaire (the second part of the questionnaire).
‡Only participants answering the general section of the questionnaire were invited to complete the food frequency questionnaire.

© 2005 Lippincott Williams & Wilkins
Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 2. Description of the Study Population and Internet Habits by Type of Questionnaire

	Printed Questionnaire (n = 188)	Web Questionnaire (n = 149)	Web Questionnaire With Feedback (n = 146)
Age, mean	38	39	39
Time spent answering the questionnaire (minutes); mean	19	18	19
Sex; %			
Women	52	54	49
Men	48	46	51
Body mass index; mean	24.7	24.0	24.7
Current daily smokers; %	10	11	10
Education; %			
Low	18	12	19
Middle	34	31	36
High	48	57	45
Breakfast habits; %			
Daily	76	76	75
4–6 times a week	9	6	10
Fewer than 4 times a week	15	18	14
Use of Internet; %			
Daily	61	70	68
Weekly	16	22	21
Monthly	7	1	3
Seldom or never	13	5	7
Used it for the first time*	N/A	1	1
Can arrange for privacy when using Internet; %			
Yes	77	87	86
No	10	7	8
Do not know	11	7	6
Use of Internet banking; %			
Yes	53	63	64
No	46	37	36

*Answering the Web questionnaire was the first time respondent used the Internet.

when using a Web questionnaire instead of a printed questionnaire. Personalized feedback in the Web questionnaire further increased the compliance rate for a second questionnaire. Total response rates for the second part of the questionnaire were similar for the printed and the Web questionnaires.

ACKNOWLEDGMENTS

We thank Netsurvey for the use of their Web-survey software and hardware.

REFERENCES

1. Baer A, Saroiu S, Koutsky L. Obtaining sensitive data through the Web: an example of design and methods. *Epidemiology*. 2002;13:640–645.
2. Netsurvey. Available at: http://www.netsurvey.info/. Accessed April 5, 2005.
3. Pitkow J, Recker M. Results from the First World-Wide Web User Survey 1994. Available at: http://www.gvu.gatech.edu/user_surveys/survey-01-1994/. Accessed April 5, 2005.
4. Kehoe C, Pitkow J, Sutton K, Aggarwal G, Rogers J. Results from the Tenth World Wide Web User Survey 1999. Available at: http://www.gvu.gatech.edu/user_surveys/survey-1998-10/tenthreport.html. Accessed April 5, 2005.

A Practical Method for Collecting 3-Day Food Records in a Large Cohort

Ann Shattuck Kolar, Ruth E. Patterson,*† Emily White,*† Marian L. Neuhouser,* Laura L. Frank,* Judi Standley,* John D. Potter,*† and Alan R. Kristal*†*

Background: Recent studies suggest that diet records are more valid measures of nutrient intake than are food-frequency question-

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

naires. However, food records are considered unsuitable for large studies due to the need to train participants and to review and correct completed records.

Methods: We evaluated a self-administered 3-day food record protocol in Washington State. One hundred men and women age 50–76 years were mailed a food record and serving-size booklet. Sixty-five people returned a completed food record and were subsequently interviewed to obtain missing information. The food records were analyzed with and without added information from the interview.

Results: The most common error was incomplete description, which affected 8% of recorded foods. Differences in mean nutrient intake between the uncorrected and corrected records were within 5%, and nutrient estimates from the 2 methods were highly correlated.

Conclusions: This streamlined protocol yielded data comparable to those collected by more burdensome protocols, suggesting that the use of food records may be feasible in large cohort studies.

Submitted 13 February 2004; final version accepted 14 March 2005.

From the *Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, WA; and the †Department of Epidemiology, School of Public Health and Community Medicine, University of Washington, Seattle, WA.

Funding for this study provided by NCI, RO1 CA74846.

Correspondence: Ann Shattuck Kolar, Fred Hutchinson Cancer Research Center, P.O. Box 19024 M3-B232, Seattle, WA 98109-1024. E-mail: akolar@fhcrc.org.

DOI: 10.1097/01.ede.0000165363.27323.ac

Many scientists have questioned the adequacy of using food-frequency questionnaires (FFQs) for assessing diet in epidemiologic studies.[1–4] Studies using doubly labeled water and other biomarkers suggest that diet records are more valid than FFQs and that the error associated with FFQs is greater than previously estimated.[5–9] However, food records have generally been considered unsuitable for use in large epidemiologic studies because of the need for in-person training and for review and correction of completed records, as well as their prohibitive costs.[10–12] Despite these limitations, 2 large cohort studies in Europe[13,14] recently used food records; 1 of these[14] reported stronger associations of diet with disease outcomes using food records compared with FFQs. These results motivate the development of methods for collecting food records that are feasible for large epidemiologic studies.

This article evaluates an innovative, mailed protocol for collecting 3-day food records using a booklet specially designed to reduce the need for in-person training and post-completion review by a nutritionist. We assessed both the completeness of the self-administered dietary record and the precision of nutrient intake measures by comparing the uncorrected records (analyzed as received) to records "corrected" by nutritionist interview.

METHODS

This pilot study was conducted among a random sample of 100 men and women in the VITamins And Lifestyle (VITAL) cohort study of western Washington State.[15] A food record, detailed instructions, and return envelope were mailed to study participants. This mailing was followed by a reminder postcard to all participants and a telephone call to nonresponders.

The 3-day food record booklet was designed to be entirely self-administered. It contained instructions for recording food intake (including how to describe food preparation methods, added fats, brand names, and ingredients of mixed dishes and recipes), as well as an example of a correctly completed day's record. The booklet also contained 12 questions on food-use patterns to collect information typically obtained during review of completed food records; these responses were used to assign default values when food records were incomplete. Participants also received a 12-page serving-size booklet containing photographs and other measurement tools to facilitate accurate quantification of foods and beverages consumed. Both instruments can be viewed at http://ffq.fhcrc.org.

Trained staff entered food records into the Nutrition Data System for Research (NDS-R) software (version 4.04, Food and Nutrient Database 32)[16,17] using a set of rules to standardize entry of foods with incomplete information. Records were additionally coded for the numbers of foods that were missing the following types of required information: serving size, food description, preparation method, and mixed food or recipe ingredients. Within 1 week of receiving completed food records, a nutritionist telephoned respondents to obtain missing and incomplete information. These corrected records were then entered into NDS-R by staff who was unaware of decisions made during analysis of the uncorrected records.

Error rates (defined as the number of omissions divided by the number of foods subject to that omission) were calculated for each type of missing information for each day's intake. To test whether error rates differed by participant characteristics, we used linear regression models with daily error rates (as response variables) weighted for the number of foods in each category of error. To assess the impact of recording errors on nutrient estimates, we calculated bias as the mean difference between uncorrected and corrected records, and we computed precision as Pearson's correlation coefficient. We used intraclass correlations among 3 days of food records to compare day-to-day variation in nutrient intake between the uncorrected and corrected records. Nutrients with skewed distributions were log-transformed before analyses.

RESULTS

Sixty-eight participants (68%) returned a completed food record; 53% responded to the initial mailing and postcard, and an additional 15% returned a completed food record after the telephone reminder. We were unable to reach 3

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

participants for review of their food record; therefore, this report is limited to the 65 participants with complete data. Participants were similar to the VITAL study cohort with regard to most demographic characteristics: approximately half were women, more than 95% were white, 85% had some college education, half were overweight or obese, and 6% were current smokers.

Table 1 gives data on the omissions and error rates in the self-administered 3-day food records. Participants reported consuming an average of 17.6 foods per day, of which 3% were missing portion sizes and 8% were incompletely described. Participants recorded an average of 2 foods per day that required information on how the food was prepared, but almost half of these lacked adequate details. An average of only 1 recipe (or mixed food) was consumed per day, and almost 40% were inadequately specified. There were no significant differences in these error rates by age, sex, education, or body mass index (BMI), but there was a trend toward fewer incomplete recipes with increasing education and more missing portion sizes with increasing BMI.

Table 2 gives comparisons of uncorrected to corrected food records. Differences in mean nutrient intake between the uncorrected and corrected records were small. The uncorrected records overestimated fat intake by 4 g (5.5%) and percent energy from fat by 1 percentage point (3.6%). The uncorrected records underestimated vitamin C by approximately 4 mg (4.8%) and calcium by 42 mg (5.4%). Correlation coefficients between nutrients estimated from the uncorrected and corrected records were very high, ranging from 0.87 to 1.00.

Intraclass correlations among the 3 days of records ranged from 0.30 to 0.77 (for β-carotene and alcohol, respectively) in the uncorrected food records and from 0.24 to 0.78 in the corrected records. The largest discrepancies were 0.47 versus 0.58 for percent energy from carbohydrate and 0.51 versus 0.59 for fat (g) for the uncorrected and corrected food records, respectively. There were no other substantial differences in the day-to-day nutrient variation between the 2 food record methods (data not shown).

DISCUSSION

This pilot study found reasonable data quality using self-administered food records. Comparing uncorrected records with corrected records, mean intake of nutrients differed by less than 6%, correlations were uniformly high, and within-person day-to-day variability in intake of most nutrients was very similar. Based on these findings, we conclude that food records analyzed without subsequent augmentation by participant interview could be a methodologically and scientifically viable option in cohort studies examining diet and chronic disease risk.

We anticipated that respondents would omit details regarding food descriptions, portion sizes, and preparation methods. The food record and serving size booklets were therefore carefully designed to instruct and prompt participants to record complete information. Differences in mean nutrient intake between uncorrected and corrected records were modest and correlations comparing nutrients from the 2 methods were high. However, our study design did not allow evaluation of whether modifications to the food record booklet decreased respondent error, and further research is needed to address this question.

The use of food records in large-scale cohort studies is often dismissed because of the expense for analysis of these records for the entire cohort. However, if a food record protocol does not require in-person instruction and review, a case–cohort or nested case–control design could be used and only records from a small subset of the cohort (ie, cases and controls) would need analysis. Our results suggest that high-quality food record data can be collected without record review, such that completed food records could be stored for many years and analyzed only as needed.

There are several limitations inherent in the design of this pilot study. As is true of all cohort studies, participants in the VITAL cohort and this pilot study were a volunteer sample characterized by their willingness to complete an extensive questionnaire. The response rate was 68%; although low response rates decrease study power, it is unlikely to introduce selection bias, because participation in a cohort study is generally not jointly affected by exposure and future (unknown) disease incidence. The sample size of this pilot study was small and could not test whether food record quality differed by demographic characteristics. Participants kept food records on 3 consecutive days, beginning on their day of choice, and 42% of records included at least 1 weekend day. Although some studies report that diets differ by day of the week,[18] it is not clear that this would affect comparisons of uncorrected to corrected records. Lastly, this study examined only 1 source of measurement error: eliminating in-person instruction and post-record review of completed food records. We did not examine errors due to having only 3 days of records, behavior change due to record keeping, or other potential inaccuracies in recording food intake.

In conclusion, this pilot study demonstrated that a self-administered protocol for collecting food records has potential for use in large cohort studies. By using a streamlined protocol and a modified version of a food record booklet, in-person training and review of food records was not essential to collect high-quality food record data. This method of collecting food records should be further improved and evaluated, and its use in cohort studies is needed to evaluate whether food records are superior to FFQs in studies of diet and disease risk.

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 1. Mean Number of Foods and Omissions per Person per Day on Self-Administered 3-Day Food Records in the VITamins and Lifestyle (VITAL) Cohort Pilot Study

		Serving Size Amount			Description of Food Item*			Preparation Method[†‡]			Recipe or Mixed Food Ingredients[‡§]		
Characteristic	No.	Mean No. of Foods	Mean No. of Omissions	%	Mean No. of Foods	Mean No. of Omissions	%	Mean No. of Foods	Mean No. of Omissions	%	Mean No. of Foods	Mean No. of Omissions	%
Overall	65	17.6	0.4	2.8	17.6	1.3	7.8	1.9	0.9	46.6	1.2	0.4	37.6
Sex													
Males	32	18.9	0.4	2.3	18.9	1.2	6.6	1.9	0.8	47.1	1.4	0.5	35.6
Females	33	16.3	0.5	3.2	16.3	1.4	9.0	1.9	0.9	46.1	1.1	0.3	39.5
Education													
≤High school	9	17.0	0.6	3.4	17.0	1.3	8.2	1.7	1.1	60.8	0.7	0.5	72.4
Some college	27	18.0	0.4	2.9	18.0	1.2	7.5	1.9	0.7	38.5	1.1	0.4	38.6
College graduate/ advanced degree	27	17.5	0.4	2.3	17.5	1.4	8.2	2.0	1.0	49.4	1.4	0.4	28.6
Body mass index (kg/m^2)													
Normal (18.5–<25)	29	17.6	0.2	1.3	17.6	1.0	6.1	2.0	0.9	47.2	1.2	0.5	45.0
Overweight (25–<30)	16	18.1	0.5	2.4	18.1	1.3	6.8	2.0	0.8	41.9	1.3	0.3	23.7
Obese (≥30)	13	17.2	0.7	4.6	17.2	1.6	11.0	1.7	0.7	44.9	1.2	0.4	43.2

*Mean number of foods requiring complete description; mean number omissions = foods with incomplete description of type of milk, cereal, cookies, and so on.
[†]Mean number of foods requiring preparation method; mean number of omissions = foods missing a preparation method.
[‡]Records with no foods in these categories were removed when calculating these error rates and statistical analyses.
[§]Mean number of recipes or mixed foods (eg casseroles, soups, or sandwiches); mean number omissions = recipes or mixed foods missing ingredients.

© 2005 Lippincott Williams & Wilkins
Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 2. Nutrient Intake From Self-Administered 3-Day Food Records, Uncorrected vs. Corrected After Review With Respondents for Completeness (n = 65)

Nutrient	Uncorrected Mean	Corrected Mean	Mean Difference (SE)	Pearson Correlation (95% CI)
Energy (kcal)	1922	1883	38.4 (20.7)	0.96 (0.94–0.98)
Fat (g)	74.9	71.0	3.9 (1.3)	0.95 (0.91–0.97)
Percent energy from fat	34.3	33.1	1.2 (0.4)	0.92 (0.88–0.95)
Carbohydrate (g)	240	238	2.1 (3.0)	0.96 (0.93–0.97)
Percent energy carbohydrate	50.6	51.2	−0.6 (0.4)	0.93 (0.88–0.95)
Protein (g)	73.1	74.0	−1.0 (1.2)	0.91 (0.85–0.94)
Percent energy from protein*	15.6	16.2	−0.6 (0.2)	0.87 (0.79–0.92)
Alcohol (g)*	6.0	6.2	−0.2 (0.2)	1.00 (0.99–1.00)
Percent energy from alcohol*	1.9	2.0	−0.1 (0.1)	0.99 (0.99–1.00)
Vitamin C (mg)	89.1	93.4	−4.3 (2.1)	0.94 (0.90–0.96)
β-carotene (μg)*	2973	3099	−126.4 (93.6)	0.92 (0.87–0.95)
Vitamin E (mg)*	7.5	7.5	0.0 (0.2)	0.96 (0.93–0.97)
Calcium (mg)	785	827	−42.0 (15.3)	0.93 (0.89–0.96)
Fiber (g)	19.8	20.5	−0.7 (0.3)	0.95 (0.93–0.97)

*Log-transformed.
SE, standard error.

ACKNOWLEDGMENTS

We thank Kayla Stratton for the statistical analyses for this report. We also gratefully acknowledge the Nutrition Assessment Shared Resource at the Fred Hutchinson Cancer Research Center for assistance in the development of the food-record and serving-size booklets and the review and coding of the food-record data.

REFERENCES

1. Freedman LS, Carroll RJ, Wax Y. Estimating the relation between dietary intake obtained from a food frequency questionnaire and true average intake. *Am J Epidemiol*. 1991;134:310–320.
2. Subar AF, Thompson FE, Kipnis V, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *Am J Epidemiol*. 2001;154: 1089–1099.
3. Byers T. Food frequency dietary assessment: how bad is good enough? *Am J Epidemiol*. 2001;154:1087–1088.
4. Dennis LK, Snetselaar LG, Nothwehr FK, et al. Developing a scoring method for evaluating dietary methodology in reviews of epidemiologic studies. *J Am Diet Assoc*. 2003;103:483–487.
5. McKeown NM, Day NE, Welch AA, et al. Use of biological markers to validate self-reported dietary intake in a random sample of the European Prospective Investigation into Cancer United Kingdom Norfolk cohort. *Am J Clin Nutr*. 2001;74:188–196.
6. Day N, McKeown N, Wong M, et al. Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. *Int J Epidemiol*. 2001;30:309–317.
7. Kipnis V, Midthune D, Freedman LS, et al. Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol*. 2001;153:394–403.
8. Subar AF, Kipnis V, Troiano RP, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN Study. *Am J Epidemiol*. 2003;158:1–13.
9. Kipnis V, Subar AF, Midthune D, et al. Structure of dietary measurement error: results of the OPEN Biomarker Study. *Am J Epidemiol*. 2003;158:14–21.
10. Armstrong BK, White E, Saracci R. *Principles of Exposure Measurement in Epidemiology. Monographs in Epidemiology and Biostatistics*, vol 21. Oxford: Oxford University Press; 1992, 1994.
11. Willett W. *Nutritional Epidemiology*, 2nd ed. New York: Oxford University Press; 1998:55.
12. Kristal AR, Shattuck AL, Williams A. *Food Frequency Questionnaires for Diet Intervention Research. 17th National Nutrient Databank Conference*. Baltimore: International Life Sciences Institute; 1994;110–125.
13. Rissanen TH, Voutilainen S, Virtanen JK, et al. Low intake of fruits, berries and vegetables is associated with excess mortality in men: the Kuopio Ischaemic Heart Disease Risk Factor (KIHD) Study. *J Nutr*. 2003;133:199–204.
14. Bingham SA, Luben R, Welch A, et al. Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet*. 2003;362:212–214.
15. White E, Patterson RE, Kristal AR, et al. The Vitamins and Lifestyle (VITAL) Cohort Study: study design and characteristics of supplement users. *Am J Epidemiol*. 2004;159:83–93.
16. Schakel SF, Sievert YA, Buzzard IM. Sources of data for developing and maintaining a nutrient database. *J Am Diet Assoc*. 1988;88:1268–1271.
17. Schakel SF, Buzzard IM, Gebhardt SE. Procedures for estimating nutrient values for food composition databases. *J Food Composition Anal*. 1997;10:102–114.
18. Thompson FE, Kipnis V, Subar AF, et al. Evaluation of two brief instruments and a food-frequency questionnaire to estimate daily number of servings of fruit and vegetables. *Am J Clin Nutr*. 2000;71:1503–1510.

Hypospadias in Sons of Women Exposed to Diethylstilbestrol In Utero

Julie R. Palmer, Lauren A. Wise,* Stanley J. Robboy,† Linda Titus-Ernstoff,‡ Kenneth L. Noller,§ Arthur L. Herbst,¶ Rebecca Troisi,‖ and Robert N. Hoover‖*

Background: Diethylstilbestrol (DES) is a synthetic estrogen that was widely prescribed to pregnant women before 1971. DES increases the risk of breast cancer in women who took the drug and the

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

risk of reproductive tract abnormalities in their offspring. Dutch investigators have reported a 20-fold increase in risk of hypospadias among sons of women who were exposed to DES in utero. We assessed this relation in data from an ongoing study of DES-exposed persons.

Methods: Several U.S. cohorts of women with documented exposure in utero to DES have been followed by mailed questionnaires since the 1970s. Comparison subjects are unexposed women of the same ages. In 1997, participants were asked about congenital abnormalities in their children. We calculated prevalence odds ratios for the risk of hypospadias in sons of exposed mothers relative to sons of unexposed mothers using generalized estimating equations to adjust for multiple sons per mother and controlling for maternal age at the son's birth.

Results: We obtained data from 3916 exposed and 1746 unexposed women. These women reported a total of 13 liveborn sons with hypospadias (10 exposed, 3 unexposed). The prevalence odds ratio for risk of hypospadias among the exposed was 1.7 (95% confidence interval = 0.4–6.8).

Conclusions: Our findings do not support a greatly increased risk of hypospadias among the sons of women exposed to DES in utero, as has been previously reported.

Submitted 29 July 2004; accepted 4 October 2004.

From the *Slone Epidemiology Center, Boston University, Boston, Massachusetts; the †Department of Pathology, Duke University Medical Center, Durham, North Carolina; the ‡Norris Cotton Cancer Center, Dartmouth Hitchcock Medical Center, Lebanon, New Hampshire; the §Department of Obstetrics and Gynecology, New England Medical Center, Boston, Massachusetts; the ¶Department of Obstetrics and Gynecology, University of Chicago, Illinois; and the ‖Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland.

Supported by National Cancer Institute contracts N01-CP-21168, N01-CP-51017, and N01-CP-01289.

Correspondence: Julie R. Palmer, Slone Epidemiology Center, Boston University, 1010 Commonwealth Avenue, Boston, MA 02215. E-mail: jpalmer@slone.bu.edu.

DOI: 10.1097/01.ede.0000164789.59728.6d

Diethylstilbestrol (DES) is a synthetic estrogen that was commonly prescribed during the 1950s and 1960s to women with high-risk pregnancies and to many women with normal pregnancies.[1] The daughters of women who took DES while pregnant have an increased prevalence of reproductive tract abnormalities leading to spontaneous abortion and infertility and a sharply increased incidence of clear cell adenocarcinoma of the vagina and cervix.[2–4] Data on the sons are less definitive, but there have been reports of an increased prevalence of epididymal cysts, hypotrophic testes, and testicular varicoceles.[5–7]

Animal studies suggest that DES may increase susceptibility to reproductive tract tumors in subsequent generations.[8–11] One possible mechanism involves genetic or epigenetic changes in the germ cells that are transmitted to future generations. Alternatively, there may be somatic cell changes that affect uteroplacental function in the DES-exposed daughter. There are almost no data on third-generation effects of DES in humans. Kaufman and Adam[12] examined 28 daughters of women with DES exposure in utero and found no occurrences of the reproductive tract abnormalities that are relatively common in second-generation exposed. However, in a recent study of Dutch women with fertility problems,[13] the prevalence of hypospadias was more than 20 times greater among sons born to women exposed prenatally to DES than among sons of unexposed women.

Hypospadias is a relatively common birth defect in which the urethral opening is located on the ventral surface of the penis or on the scrotum.[14,15] Most cases require surgical correction. The causes of hypospadias are unknown, but it may be related to levels of sex hormones during the period of penile and urethral development (weeks 8–14 of gestation).[16–18]

The Dutch study provides the first epidemiologic data suggesting a third-generation effect of DES exposure in humans. The current analysis was untaken to assess whether those results could be confirmed in data from an ongoing DES follow-up study in the United State.

METHODS

A collaborative follow-up study of women with prenatal exposure to DES and unexposed women of the same ages has been in progress since 1992.[19] The cohort was assembled from several existing cohorts: (1) women exposed to DES who, in the late 1970s, enrolled in the National Cooperative Diethylstilbestrol Adenosis Project, their unexposed sisters, and age-matched unexposed women chosen from the same hospital birth record sources as the exposed[20]; (2) DES-exposed and unexposed daughters of women who participated in a randomized clinical trial of DES at the University of Chicago in 1951–1952[21]; (3) women exposed to DES and their unexposed siblings born to women treated for infertility in a private Boston-area infertility practice[19]; and (4) DES-exposed and unexposed daughters of women who previously participated in a large study of DES health effects in the mothers.[22] In each cohort, unexposed participants were drawn from the same sources as exposed participants and were similar with respect to year of birth and socioeconomic status. In 1992, the National Cancer Institute sponsored an effort to combine the existing cohorts, trace the women originally identified as eligible, and follow participants with mailed questionnaires. One percent each of the exposed and unexposed groups had died; 18% of exposed and 16% of unexposed were lost to follow-up, which left 81% of exposed and 83% of unexposed who were successfully located and have been followed with mailed questionnaires through 1997.

We used the 1997 questionnaire data on liveborn children to determine the denominator of liveborn sons. Women were asked an open-ended question, "Were any of your children born with congenital abnormalities?" Participants were asked to record the date of birth and an explanation of

© 2005 Lippincott Williams & Wilkins
Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

the abnormality. Coding of responses was conducted without knowledge of exposure status.

Odds ratios (ORs) and 95% confidence intervals (CIs) for the association of DES exposure with prevalence of hypospadias were calculated using generalized estimating equations to account for multiple sons per mother.[23] All models controlled for maternal age at birth (<30, 30–34, 35+ years), which was associated with an increased risk of hypospadias. Maternal smoking, years of education, and history of abortion in the grandmother were not materially associated with risk of hypospadias and were not included in the final models.

RESULTS

There were 2522 liveborn sons among women exposed to DES and 1336 liveborn sons among women who were not exposed. Women in the exposed cohort reported 10 sons with hypospadias (2 from the same mother), for a prevalence of 4.0 cases per 1000 male births. Women in the unexposed cohort reported 3 sons with hypospadias, for a prevalence of 2.2 cases per 1000 male births. The OR for hypospadias for sons of exposed women relative to sons of unexposed women was 1.7 (95% CI = 0.4–6.8). The association was similar whether the prenatal exposure to DES began in the first trimester of pregnancy (1.7; 0.3–8.8) or later in pregnancy (1.9; 0.4–9.9).

We repeated our analyses among women with fertility problems (those who reported having tried to become pregnant for 12 months or more without success and who consulted a physician for that reason) and among all other women. Among women with no history of fertility problems, DES exposure was not associated with hypospadias (1.1; 0.2–4.9). Among those who reported fertility problems, it was not possible to compute an OR because there were no cases among the 145 unexposed women. The lower bound of a 95% CI for that estimate was 0.2, indicating statistical compatibility with no association. The prevalence of hypospadias among sons born to DES-exposed women with a history of infertility was 8.5/1000 male births (CI = 2.8/1000–20/1000).

DISCUSSION

Our study does not confirm the previous report of a 20-fold increase in risk of hypospadias among sons of mother exposed to DES in utero.[1] The overall prevalence OR in our study was only 1.7, with an upper 95% CI of 6.6. Results were based on similar numbers of cases (12 in the Dutch study and 13 in ours), although our study had more exposed mothers and somewhat greater statistical power (2522 exposed and 1336 unexposed sons, compared with 205 exposed and 8729 unexposed sons in the Dutch study).

The prevalence of hypospadias among liveborn males in the United States is estimated to be 2 to 6 cases per 1000.[14,15] The prevalences observed in both our exposed and unexposed cohorts fall within that range. However, comparisons with published rates are problematic because the published rates are presumably based on a more complete ascertainment of cases.

One major difference between the present study and the Dutch study is the composition of the cohorts. The Dutch study was conducted within a cohort of subfertile women,[13] raising the possibility that the findings were influenced by factors related to infertility that could not be controlled. All mothers had been unable to conceive after 1 year or more of frequent unprotected intercourse and had sought medical care at infertility centers. Approximately half had undergone in vitro fertilization. By contrast, in the present study only approximately 20% of women reported difficulty becoming pregnant. When we confined the analysis to women who reported seeing a doctor for infertility, that is, a cohort similar to the Dutch cohort, we observed a higher prevalence of hypospadias in the sons of women exposed to DES (5/590) than in sons of unexposed women (0/145), although these risks are far too imprecise to establish whether they are different.

A perhaps more important difference between the 2 studies is related to the quality of data on exposure. It is well-established that reporting of maternal DES use is poor.[24,25] In our study, DES exposure status had been confirmed by medical record review before any outcomes occurred. Thus, misclassification of exposure was unlikely and could not have been differential due to outcome status. In the Dutch study,[13] information on DES exposure was obtained after the occurrence of hypospadias. Participants who gave birth to a son with hypospadias may have been more likely to learn of or recall being told of their DES exposure. It is likely that there was considerable underreporting of DES exposure by the women who delivered healthy babies.

The small excess risk of hypospadias in our own study could be due to differential reporting of hypospadias. As in the Dutch study, we relied on the mother's report of hypospadias in response to a single open-ended question about giving birth to a child with a congenital abnormality. We did not obtain confirmation by medical record. DES-exposed women may have been more complete than unexposed women in their reporting of hypospadias malformations, although there was not an increased prevalence of other congenital malformations. Thus, although the small increase in risk among sons of exposed women may be due simply to statistical imprecision, we cannot exclude the possibility of some bias.

In summary, the present results from a large U.S. cohort of DES-exposed and unexposed women provide no support for a recent report[13] that maternal in utero exposure to DES greatly increases risk of giving birth to a son with hypospadias.

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

ACKNOWLEDGMENTS

We are extremely grateful to the many women who have participated in this long-term study.

REFERENCES

1. Noller KL, Fish CR. Diethylstillbestrol usage: its interesting past, important present, and questionable future. *Med Clin North Am*. 1974; 58:739–810.
2. Kaufman RH, Adam E, Hatch EE, et al. Continued follow-up of pregnancy outcomes in diethylstilbestrol-exposed offspring. *Obstet Gynecol*. 2000;96:483–489.
3. Palmer JR, Hatch EE, Rao S, et al. Infertility among women exposed prenatally to diethylstilbestrol. *Am J Epidemiol*. 2001;154:316–321.
4. Herbst AL, Ulfelder H, Poskanzer Dc. Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with tumor appearance in young women. *N Engl J Med.* 1971;284:878–881.
5. Gill WB, Schumacher GF, Bibbo M, et al. Association of diethylstilbestrol exposure in utero with cryptorchidism, testicular hypoplasia and semen abnormalities. *J Urol*. 1979;122:36–39.
6. Leary FJ, Resseguie LJ, Kurland LT, et al. Males exposed in utero to diethylstilbestrol. *JAMA*. 1984;252:2984–2989.
7. Wilcox A, Baird D, Weinberg C, et al. Fertility in men exposed prenatally to diethylstilbestrol. *N Engl J Med*. 1995;332:1411–1416.
8. Newbold RR, Hanson RB, Jefferson WN, et al. Proliferative lesions and reproductive tract tumors in male descendants of mice exposed developmentally to diethylstilbestrol. *Carcinogenesis*. 2000;21:1355–1363.
9. Newbold RR, Hanson RB, Jefferson WN, et al. Increased tumors but uncompromised fertility in the female descendants of mice exposed developmentally to diethylstilbestrol. *Carcinogenesis*. 1998;19:1655–1663.
10. Turusov VS, Trukhanova LS, Parfenov YuD, et al. Occurrence of tumours in the descendants of CBA male mice prenatally treated with diethylstilbestrol. *Int J Cancer.* 1992;50:131–135.
11. Walker BE. Tumors of female offspring of mice exposed prenatally to diethylstilbestrol. *J Natl Cancer Inst*. 1984;73:133–140.
12. Kaufman RH, Adam E. Findings in female offspring of women exposed in utero to diethylstilbestrol. *Obstet Gynecol*. 2002;99:197–200.
13. Klip H, Verloop J, van Gool JD, et al. Hypospadias in sons of women exposed to diethylstilbestrol in utero: a cohort study. *Lancet*. 2002;359: 1102–1107.
14. Sweet R, Schrott H, Kurland R, et al. Study of the incidence of hypospadias in Rochester, Minnesota, 1940–1970, and a case-control comparison of possible etiologic factors. *Mayo Clin Proc*. 1974;49:52–58.
15. Paulozzi LJ. International trends in rates of hypospadias and cryptorchidism. *Environ Health Perspect*. 1999;107:297–302.
16. Baskin LS. Hypospadias and urethral development. *J Urol*. 2000;163: 951–956.
17. Dolk H. Rise in prevalence of hypospadias. *Lancet*. 1998;351:770.
18. Akre O, Lipworth L, Cnattingius S, et al. Risk factor patterns for cryptorchidism and hypospadias. *Epidemiology*. 1999;10:364–369.
19. Hatch EE, Palmer JR, Titus-Ernstoff L, et al. Cancer risk in women exposed to diethylstilbestrol in utero. *JAMA*. 1998;280:630–634.
20. Labarthe D, Adam E, Noller KL, et al. Design and preliminary observations of the National Cooperative Diethylstilbestrol Adenosis (DESAD) Project. *Obstet Gynecol*. 1978;51:453–458.
21. Bibbo M, Gill WB, Azizi F, et al. Follow-up study of male and female offspring of DES-exposed mothers. *Obstet Gynecol*. 1977;49:1–8.
22. Greenberg ER, Barnes AB, Resseguie L, et al. Breast cancer in mothers given diethylstilbestrol in pregnancy. *N Engl J Med*. 1984;311:1393–1398.
23. Horton NJ, Lipsitz SR. Review of software to fit generalized estimating equation (GEE) regression models. *Am Stat*. 1999;53:160–169.
24. Tilley BC, Barnes AB, Bergstralh E, et al. A comparison of pregnancy history recall and medical records: Implications for retrospective studies. *Am J Epidemiol*. 1985;121:269–281.
25. Sharp GB, Cole P, Anderson D, et al. Clear cell adenocarcinoma of the lower-genital tract: Correlation of mother's recall of diethylstilbestrol history with obstetrical records. *Cancer*. 1990;66:2215–2220.

Smoking and *Helicobacter pylori* Infection in a Sample of U.S. Adults

Victor M. Cardenas and David Y. Graham†*

Background: Although the prevalence of *Helicobacter pylori* infection has been assessed in the National Health and Nutritional Examination Survey (NHANES), its possible relation with smoking has not been fully explored.

Methods: We used the 1999–2000 NHANES data to examine whether smoking affects the prevalence of *H. pylori* infection in adults (n = 3689). The prevalence of *H. pylori* infection, as assessed by anti-*H. pylori* IgG enzyme-linked immunoassay, was computed according to smoking history and levels of serum cotinine.

Results: Current smoking was associated with an increased prevalence of *H. pylori* infection (prevalence odds ratio =1.9; 95% confidence interval = 1.4–2.5) after controlling for possible confounders. Current smoking, as measured by increased serum cotinine, was also associated with an increased prevalence of *H. pylori* infection (1.6; 1.3–2.0). There was no evidence of a dose–response relation.

Conclusions: Smoking may contribute to the persistence of *H. pylori* infection.

Submitted 24 July 2004; final version accepted 21 March 2005.

From the *UT School of Public Health at Houston–El Paso Campus, El Paso, TX; and †Baylor College of Medicine–Gastroenterology Section, Houston, TX.

Supported in part by the Office of Research and Development, Medical Research Service, Department of Veterans Affairs and by Public Health Service grant DK56338.

Correspondence: Victor M. Cardenas, 1100 Stanton Ave., Ste. 110F, El Paso, TX 79902. E-mail: Victor.Cardenas@uth.tmc.edu.

DOI: 10.1097/01.ede.0000165365.52904.4a

Smoking and *Helicobacter pylori* infection are independent risk factors for peptic ulcer disease and gastric cancer.[1–4] Six of 9 published studies that have looked into the relation of smoking with *H. pylori* infection have found evidence consistent with a weak positive association.[5–14] For example, a previous study[11] based on the Third National Health and Nutritional Examination Survey (NHANES) data reported no association overall, but a dose–response relation between pack-years of smoking and *H. pylori* infection was found among blacks. We report a cross-sectional analysis of the current NHANES (1999–2000) that examined the relation between *H. pylori* infection and smoking as measured both through standard questionnaire and in relation to serum cotinine levels.

METHODS

We used the 1999–2000 NHANES data, for which details on survey design and methods are presented elsewhere.[15] Briefly, NHANES provides estimates of health conditions in the civilian, noninstitutionalized U.S. population.

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

H. pylori infection was assessed with a commercially available immunoglobulin G (IgG) enzyme-linked immunosorbent assay (ELISA). As determined by a previous validation study, those with an immune status ratio >0.9 were considered infected.[16] In our analysis of serum cotinine levels, we used a cotinine above 15 ng/mL to define current smoking.[17] Using serum cotinine as the gold standard, the smoking questionnaire data had a sensitivity of 96% (95% confidence interval [CI] = 94–97%) and a specificity of 97% (96–98%) for classifying current cigarette smoking. Only 2.5% of persons who reported that they were former smokers had cotinine >15 ng/mL.

We excluded from the analyses records from persons younger than 20 years of age (5085 of 9965, or 49%) and also records with missing data on the study variables (1191 of 4880, or 24%), leaving 3689 complete observations in the analytic dataset. Data were analyzed taking into consideration the complex survey design using SUDAAN (version 8.1; Research Triangle Institute, Research Triangle Park, NC). Using the 2000 Census as the standard population, we obtained the age-standardized prevalence of *H. pylori* infection in our study population and in NHANES III[11] using variance estimators described elsewhere.[18] We then computed the distribution of *H. pylori* infection according to the following smoking variables: current, former, and never cigarette smokers; number of cigarettes smoked; duration in years; pack-years; and cotinine levels as a dichotomous variable (>15 ng/mL) and as tertiles of the distribution of those at or above that cutoff.

RESULTS

The 2000 U.S. Census age-standardized prevalence of *H. pylori* infection in our study population was 32% (standard error [SE] = 1.4%), slightly lower than in NHANES III (age-standardized = 34%, SE = 1.5%) The age-standardized prevalence ratio *H. pylori* infection in NHANES III relative to current NHANES was 1.1 (95% CI = 1.0–1.2).

In NHANES III, all participants were found to have detectable levels of cotinine,[17] whereas 35% of participants in the 1999–2000 NHANES had no detectable levels of cotinine.[19] In our study, persons infected with *H. pylori* had a median serum cotinine of 0.11 ng/mL compared with 0.08 for those uninfected. Among *H. pylori* infected participants, the 75th percentile of the distribution of cotinine was 19.5 ng/mL, compared with 2.07 ng/mL among those uninfected.

Results of our analyses for each potential confounder and *H. pylori* are shown in Table 1. The prevalence of *H. pylori* infection increased with age, among persons of ethnic/racial groups other than non-Hispanic whites, and among foreign-born participants. There was an inverse relationship of infection prevalence with education, vitamin C intake, and alcohol intake.

As shown in Table 2, current smokers had an increased prevalence of *H. pylori* infection compared with never-smokers, with the prevalence among former smokers intermediate between the levels of current and never-smokers. After we controlled for several variables, current and former cigarette smokers had a higher prevalence of *H. pylori* infection (prevalence odds ratio [POR] = 1.9 and 1.3, respectively) compared with never-smokers. Analyses restricted to U.S.-born non-Hispanic whites (n = 1629) resulted in virtually the same point estimates: POR = 2.0 for current and POR = 1.4 for former cigarette smokers.

Table 2 also displays results for several other smoking variables. We did not find evidence of a dose–response relation between the prevalence of *H. pylori* infection and the number of cigarettes smoked, years of smoking, time since quitting smoking, pack-years of cigarette smoking, or tertiles of the distribution of serum cotinine. We found that the prevalence of *H. pylori* infection was not higher among those individuals grouped in the categories of heavier tobacco use compared with those in intermediate categories.

Study participants who had levels of serum cotinine compatible with current smoking status had an increased prevalence of infection (1.6; 1.3–2.0) compared with those classified as nonsmokers, controlling for all relevant measured covariates.

DISCUSSION

We report an association between smoking and *H. pylori* infection in a national sample of the U.S. adult population using both questionnaire data and serum cotinine. Our study is subject to several limitations. First, antibody against *H. pylori* identifies both active and past infections such that not all persons with antibody had current infection. Second, because of the cross-sectional nature of the study, we could not identify changes in infection status over time. Longitudinal studies in children and in adults show a loss of the infection over time.[20,21] The mechanism of loss of infection remains unclear, with one possibility being a bystander effect of the use of antibiotics for other indications. Information on *H. pylori* treatment was not collected in current NHANES.

The 2 general ways by which the pool of *H. pylori*-infected and ever-smoking individuals can be enriched are through the presence of an increased susceptibility or exposure (higher rates of acquisition) or by decreased loss of the infection over time. Most *H. pylori* infections are acquired in childhood, and so the risk of acquiring infection is therefore unlikely to be affected by smoking status itself. However, smoking status might reflect socioeconomic status (SES), a variable known to relate to *H. pylori* prevalence. Data against this possibility include the persistence of the association among U.S.-born non-Hispanic white participants.

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 1. Weighted Prevalence* of *H. pylori* Infection by Selected Risk Factors Among Adults (20+ Years) Participating in the Current NHANES: United States 1999–2000 (n = 3689)

	No. of Persons	Infected (%)	Multivariate Prevalence Odds Ratio†	(95% CI)
Age (years)				
20–29‡	644	22	1.0	
30–39	641	27	1.3	(0.9–1.9)
40–49	602	29	1.8	(1.2–2.5)
50–59	475	37	2.8	(2.1–3.8)
60+	1327	42	3.9	(2.6–5.8)
Sex				
Men	1738	31	0.9	(0.7–1.3)
Women‡	1951	31	1.0	
Ethnicity				
Non-Hispanic white‡	1713	22	1.0	
Non-Hispanic black	634	53	4.1	(3.1–5.5)
Mexican-American	1004	63	4.2	(2.8–6.4)
Other Hispanic	224	57	2.8	(1.6–4.8)
Other	114	46	2.5	(1.7–3.8)
Place of birth				
Outside U.S.	951	61	2.9	(2.0–4.1)
U.S.‡	2738	26	1.0	
Education				
<High school	1389	56	2.6	(2.0–3.5)
High school	832	32	1.7	(1.3–2.2)
>High school‡	1468	20	1.0	
Vitamin C intake (quintiles)				
(Lowest) 1st‡	729	38	1.0	
2nd	735	32	0.7	(0.6–1.0)
3rd	759	28	0.7	(0.5–0.9)
4th	739	29	0.6	(0.4–0.9)
(Highest) 5th	727	29	0.6	(0.5–0.9)
≥12 drinks of alcohol per year				
Yes	2450	29	0.9	(0.7–1.2)
No‡	1239	36	1.0	

*Weighted by the inverse of the probabilities of each person in the sample.
†Each measure of effect is adjusted for the remaining variables in the table.
‡Reference category.

The prevalence of *H. pylori* is related to the SES of their families during the time they were children.[22] Thus, current smoking may also reflect a lower SES in childhood. We performed a sensitivity analysis of confounding by childhood SES (ie, externally adjusted[23]) using prevalence estimates of regular smoking by type of father's job (ie, manual and nonmanual)[24] and of the association between low childhood SES and *H. pylori* infection.[22] Even if we reduce our estimate of the association between smoking and *H. pylori* by 30% to 40% (as suggested by the sensitivity analysis), smokers still had an excess prevalence of *H. pylori* infection. However, SES at early ages strongly predicts SES at later ages; adjustment by education should have removed most confounding by childhood SES.[25]

Conflicting results of this and prior studies may reflect different criterion validity, and considerable variation in both the prevalence of *H. pylori* infection (21–66%) and current smoking (20–40%) across study populations.[5–14] We found

© 2005 Lippincott Williams & Wilkins
Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 2. Weighted* Prevalence of *H. pylori* Infection According to Cigarette Smoking Among Adults (20+ Years) Participating in the Current NHANES: United States 1999–2000

	No. of Persons	Infected (%)	Multivariate† Prevalence Odds Ratio	(95% CI)
Cigarette smoking status				
Current	774	38	1.9	(1.4–2.5)
Former	1010	31	1.3	(1.0–1.7)
Never	1853	28	1.0	
Amount of smoking (cigarettes per day; current smokers)§				
1–10	285	48	2.2	(1.5–3.2)
≥Half a pack	353	34	1.8	(1.3–2.4)
Duration of smoking (years; current smokers)¶				
≤15	567	31	1.8	(1.3–2.4)
16–31	543	33	1.4	(1.0–1.9)
32+	547	44	1.5	(1.1–2.1)
Time since quitting (years; never-smokers)				
<1	109	25	0.9	(0.6–1.5)
1–9	261	26	1.0	(0.6–1.6)
10–20	328	37	1.5	(0.9–2.5)
>20	312	33	1.3	(0.9–1.9)
Pack-years (current and former smokers)‖				
<6	524	35	1.7	(1.3–2.3)
6–24	482	36	1.8	(1.3–2.6)
>24	521	33	1.2	(1.0–1.6)
Serum cotinine (ng/mL)				
Nonsmokers ≤15‡	2810	30	1.0	
Smokers >15	879	37	1.6	(1.3–2.0)
Tertiles				
1st (15.1–128.98)	289	36	1.6	(1.1–2.4)
2nd (129.0–234.74)	294	37	1.6	(1.2–2.2)
3rd (234.75–989.00)	296	38	1.6	(1.1–2.3)

*Weighted by the inverse of the probabilities of each person in the sample.

†All estimates are simultaneously adjusted for age, gender, ethnicity, place of birth, vitamin C diet intake, and regular alcohol consumption.

‡For all comparisons except serum cotinine, never-smokers are the reference category. For cotinine, those with cotinine levels ≤15 are the referent category.

§One hundred thirty-six current smokers did not provide information on number of cigarettes smoked.

¶One hundred twenty-seven participants with missing data to compute years smoked.

‖Two hundred fifty-seven former or current smokers with missing data to compute pack-years smoked.

correlation between strength of the association and prevalence of smoking or infection across those studies. However, never-smokers in NHANES III had a 6% higher prevalence of *H. pylori* infection than in our study population, which might explain the difference. We did not find a dose–response association between *H. pylori* with amount or duration of cigarette smoking. This may be because a dose–response relation is nonlinear, because the effect is dose-independent, or because there is no causal connection at all.

The association is biologically plausible; smokers also differ physiologically from nonsmokers in ways that might affect prevalence of infection. At least one study has shown an increased inflammatory reaction to *H. pylori* infection among smokers.[26] In addition, smokers have increased gas-

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

tric acid secretion, and smoking also inhibits both duodenal and pancreatic bicarbonate secretion, which increases the likelihood of developing duodenal ulcer disease.[27] Smokers could have an excess of persistent *H. pylori* infections due to the adverse effects of smoking on the immune system.[28] However, we hypothesize that because *H. pylori* infections within a given population are decreasing over time, the increased prevalence of *H. pylori* infection among smokers may reflect the effect of smoking on enhancing persistence of the infection.[29] For example, antibiotic therapy with single, dual, or triple drug therapy is markedly less effective among smokers than nonsmokers[30] such that on average, smokers would be less likely to have bystander eradication than nonsmokers.

Complete data from current NHANES, including data on *cagA* testing, are not yet available; additional analysis, including parallel analysis of NHANES III data, would be of interest. Further studies are warranted to examine the effect of smoking on acquiring infection, as well as on persistence of *H. pylori* infection.

ACKNOWLEDGMENTS

We thank Maureen Sanderson, Zuber Mulla, and John Bristol for helpful comments, and Melchor Ortiz for reviewing the database programming.

REFERENCES

1. Graham DY, Rakel RE, Fendrick AM, et al. Scope and consequences of peptic ulcer disease. How important is asymptomatic *Helicobacter pylori* infection? *Postgrad Med*. 1999;105:100–102.
2. Parasher G, Eastwood GL. Smoking and peptic ulcer in the *Helicobacter pylori* era. *Eur J Gastroenterol Hepatol*. 2000;12:843–853.
3. Sepulveda AR, Graham DY. Role of *Helicobacter pylori* in gastric carcinogenesis. *Gastroenterol Clin North Am*. 2002;31:517–535.
4. Tredaniel J, Boffetta P, Buiatti E, et al. Tobacco smoking and gastric cancer: review and meta-analysis. *Int J Cancer*. 1997;72:565–573.
5. Brown LM, Thomas TL, Ma JL, et al. *Helicobacter pylori* infection in rural China: demographic, lifestyle and environmental factors. *Int J Epidemiol*. 2002;31:638–646.
6. Murray LJ, Lane AJ, Harvey IM, et al. Inverse relationship between alcohol consumption and active *Helicobacter pylori* infection: the Bristol *Helicobacter* project. *Am J Gastroenterol*. 2002;97:2750–2755.
7. Brenner H, Rothenbacher D, Bode G, et al. Relation of smoking and alcohol and coffee consumption to active *Helicobacter pylori* infection: cross sectional study. *BMJ*. 1997;315:1489–1492.
8. Woodward M, Tunstall-Pedoe H, McColl K. An investigation into factors associated with *Helicobacter pylori* infection. *J Clin Epidemiol*. 2000;53:175–181.
9. Rosenstock SJ, Jorgensen T, Andersen LP, et al. Association of *Helicobacter pylori* infection with lifestyle, chronic disease, body-indices, and age at menarche in Danish adults. *Scand J Public Health*. 2000;28:32–40.
10. Murray LJ, McCrum EE, Evans AE, et al. Epidemiology of Helicobacter pylori infection among 4,742 randomly selected subjects from Northern Ireland. *Int J Epidemiol*. 1997;26:880–887.
11. Everhart JE, Kruszon-Moran D, Perez-Perez GI, et al. Seroprevalence and ethnic differences in *Helicobacter pylori* infection among adults in the United States. *J Infect Dis*. 2000;181:1359–1363.
12. Simon JA, Hudes ES, Perez-Perez GI. Relation of serum ascorbic acid to *Helicobacter pylori* serology in US adults: the Third National Health and Nutrition Examination Survey. *J Am Coll Nutr*. 2003;22:283–289.
13. Tatemachi M, Kabuto M, Tsugane S. Effect of smoking on serum pepsinogen I level depends on serological status of. *Helicobacter pylori*. *Jpn J Cancer Res*. 2001;92:243–248.
14. Moayyedi P, Axon AT, Feltbower R, et al., Leeds HELP Study Group. Relation of adult lifestyle and socioeconomic factors to the prevalence of *Helicobacter pylori* infection. *Int J Epidemiol*. 2002;31:624–631.
15. National Center for Health Statistics. National Health and Nutrition Examination Survey 1999–2000 data files. Available at: http://www.cdc.gov/nchs/about/major/nhanes/NHANES99_00.htm. Accessed March 15, 2005.
16. Everhart J, Kruszon-Moran D, Perez-Perez G. Reliability of *Helicobacter pylori* and *Cag*A serological assays. *Clin Diagn Lab Immunol*. 2002;9:412–416.
17. Pirkle JL, Flegal KM, Bernert JT, et al. Exposure of the US population to environmental tobacco smoke: the Third National Health and Nutrition Examination Survey, 1988 to 1991. *JAMA*. 1996;275:1233–1240.
18. Greenland S, Rothman KJ. Introduction to stratified analysis. In: Rothman, KJ, Greenland S, eds. *Modern Epidemiology*. Philadelphia: Lippincott-Raven; 1998:262–263.
19. DHHS CDC. Second National Report on Human Exposure to Environmental Chemicals. January 2003;79–80.
20. Malaty HM, Graham DY, Wattigney WA, et al. Natural history of *Helicobacter pylori* infection in childhood: 12-year follow-up cohort study in a biracial community. *Clin Infect Dis*. 1999;28:279–282.
21. Malaty HM, El-Kasabany A, Graham DY, et al. Age at acquisition of *Helicobacter pylori* infection: a follow-up study from infancy to adulthood. *Lancet*. 2002;359:931–935.
22. Malaty HM, Graham DY. Importance of childhood socioeconomic-status on the current prevalence of *Helicobacter pylori* infection. *Gut*. 1994;35:742–745.
23. Greenland S. Basic methods for sensitivity analysis and external adjustment. In: Rothman, KJ, Greenland S, eds. *Modern Epidemiology*. Philadelphia: Lippincott-Raven; 1998:344–347.
24. Gilman SE, Abrams DB, Buka SL. Socioeconomic status over the life course and stages of cigarette use: initiation, regular use, and cessation. *J Epidemiol Community Health*. 2003;57:802–808.
25. Power C, Hertzman C. Health, well-being and coping skills. In: Keating D, Hertzman C, eds. *Developmental Health and the Wealth of Nations: Social, Biological and Educational Dynamics*. New York: Guilford Press; 1999.
26. Shimoyama T, Everett SM, Fukuda S, et al. Influence of smoking and alcohol on gastric chemokine mRNA expression in patients with *Helicobacter pylori* infection. *J Clin Pathol*. 2001;54:332–334.
27. Dore MP, Graham DY. Pathogenesis of duodenal ulcer disease: the rest of the story. *Baillieres Best Pract Res Clin Gastroenterol*. 2000;14:97–107.
28. Johnson JD, Houchens DP, Kluwe WM, et al. Effects of mainstream and environmental tobacco smoke on the immune system in animals and humans: a review. *Crit Rev Toxicol*. 1990;20:369–365.
29. Kumagai T, Malaty HM, Graham DY, et al. Acquisition versus loss of *Helicobacter pylori* infection in Japan: results from an 8-year birth cohort study. *J Infect Dis*. 1998;178:717–721.
30. Broutet N, Tchamgoue S, Pereira E, et al. Risk factors for failure of *Helicobacter pylori* therapy: results of an individual data analysis of 2,751 patients. *Aliment Pharmacol Ther*. 2003;17:99–109.

© 2005 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.