

## **Author's response to reviews**

**Title:** Respiratory symptoms in relation to residential coal burning and environmental tobacco smoke among early adolescents in Wuhan, China: a cross-sectional study

### **Authors:**

Paivi M Salo ([salo1@niehs.nih.gov](mailto:salo1@niehs.nih.gov))  
Jiang Xia ([xiajiang@public.wh.hb.cn](mailto:xiajiang@public.wh.hb.cn))  
C. ANDERSON Johnson ([carljohn@hsc.usc.edu](mailto:carljohn@hsc.usc.edu))  
Yan Li ([yanli@public.wh.hb.cn](mailto:yanli@public.wh.hb.cn))  
Grace E Kissling ([kissling@niehs.nih.gov](mailto:kissling@niehs.nih.gov))  
Edward L Avol ([avol@hsc.usc.edu](mailto:avol@hsc.usc.edu))  
Chunhong Liu ([liups@public.wh.hb.cn](mailto:liups@public.wh.hb.cn))  
Stephanie J London ([london2@niehs.nih.gov](mailto:london2@niehs.nih.gov))

**Version: 3 Date:** 10 November 2004

**Author's response to reviews:** see over

## Comments of Dr. David Christiani and our responses

- 1) In response to a question about a possible confounding by ambient air pollution due to different areas, the authors stated that they "have run random effects models in order to adjust for the type of community in which the school is located". In the statistical analysis section, no other models but logistic model were indicated. It seemed that the authors simply added a variable containing 22 districts into the logistic models. This is not a sensible way to deal with the issue. For one thing, a logistic model is not a random effect model. For another, "district" is a factor that may have an important confounding effect on the outcomes. A better solution perhaps is to treat the "district" as a categorical variable (with a few dummy variables) in logistic models. By doing so, the effects from the different areas can be ascertained and adjusted for (if the effects do exist). Otherwise, other analytical approaches need to be considered.

*In response to Dr. Chapman's (Reviewer #1) suggestion, we ran supplementary models using Proc Genmod procedure in SAS. The procedure Proc Genmod covers a wide suite of generalized linear models including logistic models. The procedure can also accommodate the analysis of correlated data, such as one might expect among children within the same district (SAS OnlineDoc, Version 8, <http://v8doc.sas.com/sashtml/>; McCullagh P, Nelder JA: *Generalized Linear Models*. Cambridge, UK: University Press; 1989).*

*In our supplementary logistic models, we included living area using CLASS and REPEATED statements within the Genmod procedure. We used two separate approaches to include living area in the models: 1) including living area as the original three-category variable (urban, suburban, rural), and 2) including living area as the multilevel variable of school districts (22 districts). By including living area in the CLASS statement, a set of dummy variables are generated by Genmod, so that we are not simply adding a variable to the model that has 3 or 22 levels. Although the results differed only marginally, we prefer to report results from the latter approach, which better captures the community level variability in the data. An additional advantage, as noted by Dr. Chapman is that having modeling these community levels terms in this way gives control for community level variation in outdoor air pollution.*

*Please, see our revised text in the Statistical Analysis section (p. 5-6).*

- 2) The model fitting is questionable. As the authors stated, several candidate factors were considered that possibly affected or modified the relationship being studied, but only gender and school district were included in the final models. The authors claimed that inclusion of other variables did not appreciably change the associations. However, as we we read on p.9, "we did not find strong evidence suggesting difference in effect by gender" It is bewildering why the term of gender was then included in the models, but others were not, if they all had little effect on the outcomes.

*Attempts to control confounding by simply entering all possible confounders in the model may lead to more bias than they remove (Greenland S: **Modeling and Variable Selection in Epidemiologic Analysis**. Am J Public Health 1989, 79:340-349.) In our model selection process, we evaluated whether each covariate truly confounded the odds ratios (change  $\geq 10\%$ ). We did not include covariates other than gender, school district, and primary exposures of interest (coal use and ETS) in our models because inclusion of the other potential confounders did not appreciably change the odds ratios. Gender and age are basic variables that readers expect to see adjusted for because disease rates tend to differ by these variables. Therefore, we adjusted for gender, although it was not a strong confounder in this population. Age did not vary in our analyses; otherwise, we would have included that even if it had not been a confounder.*

*Because Dr. Christiani seems to feel so strongly about this issue, in the end of this response section (Table A and Table B), we present results from the analyses where we adjusted for all potential confounders (mentioned in the text). These results show that the effect to the associations is marginal; the odds ratios change very little and the magnitude of the effects remains the same. However, this approach leads to loss of 250 observations; only 4280 subjects remain in the analyses because of missing data on exposure variables. Deleting these subjects simply to adjust for "confounding" by variables that are not actually confounders does not seem justified. Therefore, we prefer presenting the results as they were in our previously revised manuscript.*

*Prior to the assessment of confounding, we had also examined potential effect modifiers. The sentence on page 9 that Dr. Christiani mentions above "We did not find strong evidence suggesting difference of effect by gender ..." refers to effect modification, not to the assessment of confounding. Indeed, if there had been strong evidence of effect modification by gender, it would not have made sense to adjust for gender as a confounder. To clarify the meaning of the sentence, we have reworded it. Please, see the revised sentence on page 9.*

*In our population, the effect of ETS on persistent cough without colds was more pronounced among boys than girls. Although we did not find strong evidence for effect modification, we included discussion on sex-specific effects. Gender differences in airway behavior, which may relate to biological and sociocultural factors, occur throughout the life span (Becklake MR, Kauffmann F. **Gender differences in airway behaviour over the***

*human life span. Thorax 1999, 54:1119-1138). In particular, vulnerability towards ETS exposure may differ by gender (Gold DR, et al. Effects of cigarette smoking on lung function in adolescent boys and girls. N Engl J Med 1996, 335:931-937; Svanes C, et al. Parental smoking in childhood and adult obstructive lung disease: results from the European Community Respiratory Health Survey. Thorax 2004, 59:295-302). In this manuscript, we present our results adjusted for gender because the evidence for modifying effects by gender remained weak.*

*Throughout the analyses, our modeling strategy followed general modeling guidelines recommended for logistic regression (Kleinbaum DG: Logistic regression. New York: Springer-Verlag; 1994).*

#### **Additional tables:**

Table A. Respiratory symptoms in relation to residential coal burning.

In this table, odds ratios (OR) are adjusted for gender, ETS, presence of pets, presence of pests, crowding, presence of older siblings, parental asthma, physical activity, living area (school district), and time spent indoors and outdoors.

Table B. Respiratory symptoms in relation to living with smokers.

In this table, odds ratios (OR) are adjusted for gender, coal use, presence of pets, presence of pests, crowding, presence of older siblings, parental asthma, physical activity, living area (school district) and time spent indoors and outdoors.

**Table A. Respiratory symptoms in relation to residential coal burning (Table 2.)**

<b>Exposure</b>	<b>Cough with colds</b>		<b>Cough without colds</b>	
	No	Yes	No	Yes
	OR*	(95% CI)	OR*	(95% CI)
Total	3231	1049	4090	190
<b>Coal use</b>				
No	1.00		1.00	
Yes	0.96	(0.81,1.14)	0.96	(0.70,1.31)
Heating	1.03	(0.81,1.31)	0.99	(0.60,1.62)
Cooking	0.83	(0.70,0.99)	0.97	(0.66,1.43)
Both	1.28	(0.99,1.66)	0.94	(0.61,1.44)
	<b>Phlegm with colds</b>		<b>Phlegm without colds</b>	
	No	Yes	No	Yes
	OR*	(95% CI)	OR*	(95% CI)
Total	3570	710	4041	239
<b>Coal use</b>				
No	1.00		1.00	
Yes	1.06	(0.94,1.19)	0.96	(0.71,1.29)
Heating	1.08	(0.89,1.32)	1.04	(0.63,1.73)
Cooking	0.93	(0.79,1.09)	0.87	(0.59,1.28)
Both	1.39	(1.08,1.78)	1.17	(0.81,1.67)
	<b>Wheeze with colds</b>		<b>Wheeze without colds</b>	
	No	Yes	No	Yes
	OR*	(95% CI)	OR*	(95% CI)
Total	3450	830	3987	293
<b>Coal use</b>				
No	1.00		1.00	
Yes	1.45	(1.14,1.84)	1.49	(1.15,1.95)
Heating	0.99	(0.68,1.44)	1.48	(0.98,2.23)
Cooking	1.57	(1.08,2.27)	1.47	(1.10,1.96)
Both	1.59	(1.03,2.45)	1.54	(0.99,2.42)

\* Odds ratios (OR) adjusted for gender, ETS, presence of pets, presence of pests, crowding, presence of older siblings, parental asthma, physical activity, school district, and time spent indoors and outdoors. Dichotomous and multilevel odds ratios are computed in separate models.

**Table B. Respiratory symptoms in relation to living with smokers (Table 3.)**

<b>Exposure</b>	<b>Cough with colds</b>		<b>Cough without colds</b>	
	No	Yes	No	Yes
	OR*	(95% CI)	OR*	(95% CI)
Total	3231	1049	4090	190
<b>Smokers in the home</b>				
No	1.00		1.00	
Yes	1.28	(1.04,1.57)	1.12	(0.78,1.60)
1 smoker	1.25	(1.02,1.54)	1.04	(0.71,1.53)
2+ smokers	1.49	(1.12,1.98)	1.55	(0.94,2.57)
	<b>Phlegm with colds</b>		<b>Phlegm without colds</b>	
	No	Yes	No	Yes
	OR*	(95% CI)	OR*	(95% CI)
Total	3570	710	4041	239
<b>Smokers in the home</b>				
No	1.00		1.00	
Yes	1.31	(1.14,1.50)	1.69	(1.15,2.49)
1 smoker	1.32	(1.14,1.52)	1.59	(1.07,2.35)
2+ smokers	1.31	(0.97,1.75)	2.35	(1.37,4.02)
	<b>Wheeze with colds</b>		<b>Wheeze without colds</b>	
	No	Yes	No	Yes
	OR*	(95% CI)	OR*	(95% CI)
Total	3450	830	3987	293
<b>Smokers in the home</b>				
No	1.00		1.00	
Yes	1.11	(0.95,1.29)	0.89	(0.69,1.15)
1 smoker	1.11	(0.96,1.30)	0.93	(0.72,1.21)
2+ smokers	1.06	(0.77,1.45)	0.68	(0.39,1.19)

\* Odds ratios (OR) adjusted for gender, coal use, presence of pets, presence of pests, crowding, presence of older siblings, parental asthma, physical activity, school district, and time spent indoors and outdoors. Dichotomous and multilevel odds ratios are computed in separate models.