Indoor air pollution from biomass combustion and acute respiratory infections in Kenya: an exposure-response study

Majid Ezzati, Daniel M Kammen

Summary

Background Acute respiratory infections (ARI) are the leading cause of the global burden of disease and have been causally linked with exposure to pollutants from domestic biomass fuels in less-developed countries. We used longitudinal health data coupled with detailed monitoring of personal exposure from more than 2 years of field measurements in rural Kenya to estimate the exposure-response relation for particulates smaller than 10 μm in diameter (PM₁₀) generated from biomass combustion.

Methods 55 randomly-selected households (including 93 infants and children, 229 individuals between 5 and 49 years of age, and 23 aged 50 or older) in central Kenya were followed up for more than 2 years. Longitudinal data on ARI and acute lower respiratory infections (ALRI) were recorded at weekly clinical examinations. Exposure to PM₁₀ was monitored by measurement of PM₁₀ emission concentration and time-activity budgets.

Findings With the best estimate of the exposure-response relation, we found that ARI and ALRI are increasing concave functions of average daily exposure to PM₁₀, with the rate of increase declining for exposures above about 1000–2000 μg/m³. After we had included high-intensity exposure episodes, sex was no longer a significant predictor of ARI and ALRI.

Interpretation The benefits of reduced exposure to PM₁₀ are larger for average exposure less than about 1000–2000 μg/m³. Our findings have important consequences for international public-health policies, energy and combustion research, and technology transfer efforts that affect more than 2 billion people worldwide.

Lancet 2001; 358: 619–24

Introduction

Acute respiratory infections (ARI) are the leading cause of the global burden of disease, accounting for more than 6% of worldwide disease and mortality, mostly in less-developed countries. Between 1997 and 1999, acute lower respiratory infections (ALRI) caused 3.5–4.0 million deaths worldwide—more than that caused by any other infectious disease. Exposure to indoor air pollution, especially to particulate matter, from combustion of biofuels (wood, charcoal, agricultural residues, and dung), has been associated with respiratory disease in less-developed countries. More than 2 billion people use biomass as their main source of domestic energy; hence international development and public-health organisations have sought to implement preventive measures to reduce this exposure. The exposure-response relation between indoor air pollution from biomass combustion and ARI is important for assessment of the benefits and effectiveness of preventive measures such as design and dissemination of improved stoves and fuels.

Health impacts of outdoor (ambient) particulate matter in urban areas of industrialised countries have been identified and quantified in epidemiological and physiological studies in the past two decades. However, these results are applicable to a small range of exposures, mostly less than 200 μg/m³, which are mainly of concern in industrialised countries. The exposure-response relation at concentrations of hundreds or thousands of μg/m³, which are typical of indoor environments in less developed countries, is relatively unknown. Because around 80% of total global exposure to particulate matter occurs indoors in developing nations, we need to assess this relation at exposure levels typical of these settings.

Research on the health effects of indoor air pollution in less-developed countries has been hindered by lack of detailed data about exposure and illness outcomes. In many epidemiological studies, researchers have used indirect measures of personal exposure, such as fuel or housing type. Because biomass fuels are very common in rural areas, in this indirect approach to exposure estimation many people are grouped into a single exposure category. Recent findings on large variations in emissions from individual stove types, and in exposure profiles within individual households, suggest that such indirect analysis and grouping of individuals reduces the reliability of estimation of exposure-response relations.

We directly examined the exposure-response relation for indoor air pollution and ARI in developing countries in a field study in rural Kenya. Our work was part of a long-term study of the relation between energy technology, indoor air pollution, and public health. We simultaneously monitored exposure to indoor air pollution and health status of individuals for more than 2 years to quantify the exposure-response relation for health and indoor particulate matter for a continuous range of exposure concentrations.
Methods
Participants
The study took place at Mpala Ranch and Research Centre, in Laikipia District, central Kenya. Cattle herding and domestic labour are the main occupations of most people who live in the 80–100 households (about 500 individuals) on the ranch; the rest work as maintenance staff. Firewood and charcoal were the main fuels used by the study group (consisting of all households on the ranch). Study households had similar tribal backgrounds (Turkana and Samburu), economic status, and diet. Houses in cattle-herding and maintenance villages were cylindrical with conical straw roofs. Stoves were unvented, and burnt firewood or charcoal (and kerosene in three or four houses). We have presented complete data on housing and energy technology elsewhere.11,24 Our research was approved by the institutional review panel for human subjects of the University Research Board, Princeton University, USA (Case number 1890) and by the Government of Kenya, under the Office of the President Research Permit No OP/13/001/25C 167. We obtained oral informed consent from all participants.

Exposure
We did research at Mpala Ranch from 1996 to 1999, including 6–8 months of background data, which included detailed demographic data for all the households and surveys of energy use, energy technology, diet, and smoking. We monitored indoor air pollution continuously and in real time (particles smaller than 10 μm in diameter [PM10] and carbon monoxide) in 55 houses, for more than 200 days, for 14–15 h per day. These households were randomly selected from those whose inhabitants had lived on the ranch for a large proportion of the study period, and from different villages and different fuel types used. We measured pollution in actual conditions of use and simultaneously recorded the location and activities of all the household members—especially activities such as cooking that would affect exposure to pollution. We also measured dispersion of pollution inside the house. We interviewed household members and local informants about household energy technology, and their activities and time budget (time-activity budget).

We estimated profiles of exposure for every individual, which accounted for daily and day-to-day variability of exposure, time-activity budget, and indoor dispersion of pollution. Our measurement and data analysis methods for estimation of personal exposure have been described.24 We divided activities into categories: cooking, non-cooking household tasks, working outside, warming around the stove, playing, resting and eating, and sleeping. We categorised outdoors as a microenvironment and divided the inside of the house into six microenvironments. For example, people could cook directly above the fire or slightly further away, and play or rest indoors or outside. Daily exposures were calculated with equation 1 (panel).

Stove emissions varied greatly throughout the day,11,24 and included intense peaks of short duration. Some household members were always closest to the fire when pollution was high, which occurred when fuel was added or moved, the stove was lit, the cooking pot was placed on or removed from the fire, or food was stirred (figure 1). Hence, mean daily concentration alone is not a sufficient measure of exposure. Therefore, in addition to mean daily concentration (m) we used two further descriptive statistics for human exposure (ie, to calculate c in equation 1): mean above the 75th percentile (m75) to adjust for household members who were always closest to the stove when pollution was high; and mean below the 95th percentile (m95) to remove the effect of short and intense pollution peaks for individuals who were absent at these times.

We chose c75 in equation 1, from m75, m, and m95 according to criteria we previously defined.24 For example: to describe people cooking very close to the stove when emissions were highest, c75 is m75 of the burning period; for people sleeping at night, when the stove was smouldering and undisturbed, c75 is m95 of the smouldering period.

Furthermore, exposure to emissions varied with between-day variation in: moisture content or density of fuel; air flow; type of food cooked; and choice of stove and fuel if the household used more than one stove or fuel. Activity patterns can also vary with seasonal changes in work and school, illness, market days, &c. Hence, we based our exposure calculations on measurements from more than 1 day. We assigned households to pollution concentration categories for calculation of m75, m95, and m, for burning and smouldering periods. We also grouped time budgets

Daily exposure relation (equation 1)

\[ E = \sum_{j=1}^{n} w_i c_j \]

where:
- \( E \) = daily exposure
- \( w_i \) = conversion (dilution) factor for the \( j \)th microenvironment
- \( c_j \) = emission concentration in the \( j \)th period of day
- \( n \) = number of activities for each person
- \( j \) = period of day—each period corresponds to one type of activity.

In equation 1: mean above the 75th percentile (\( m_{75} \)) to adjust for household members who were always closest to the stove when pollution was high; and mean below the 95th percentile (\( m_{95} \)) to remove the effect of short and intense pollution peaks for individuals who were absent at these times.

We chose \( c_{75} \) in equation 1, from \( m_{75}, m, \) and \( m_{95} \) according to criteria we previously defined.24 For example: to describe people cooking very close to the stove when emissions were highest, \( c_{75} \) is \( m_{75} \) of the burning period; for people sleeping at night, when the stove was smouldering and undisturbed, \( c_{75} \) is \( m_{95} \) of the smouldering period.

Furthermore, exposure to emissions varied with between-day variation in: moisture content or density of fuel; air flow; type of food cooked; and choice of stove and fuel if the household used more than one stove or fuel. Activity patterns can also vary with seasonal changes in work and school, illness, market days, &c. Hence, we based our exposure calculations on measurements from more than 1 day. We assigned households to pollution concentration categories for calculation of \( m_{75}, m_{95} \), and \( m \) for burning and smouldering periods. We also grouped time budgets
This model is described by $y = \mathbf{X} \cdot \beta + \mathbf{u}$: $y$=vector of illness rates for all individuals; $\mathbf{X}$=matrix of explanatory and control variables; $\beta$=vector of coefficients; and $\mathbf{u}$=independent vector of normally distributed errors (residuals). Second, we used a logistic probability model $y = F(\mathbf{X} \cdot \beta + \mathbf{u})$: $y$, $\mathbf{X}$, and $\beta$ are defined as in the linear model; $F$=cumulative logistic distribution, $P(z) = \exp(z)$ divided by $[1 + \exp(z)]$. In logit or logistic regression models the left-hand side is the probability of an event $y$ (such as illness), or $P(y)$. Because we defined our primary outcome as fraction of time with illness (ie, equivalent to rate or probability of illness), the left-hand side is only $y$. A b-logit regression with maximum-likelihood parameter estimation was used to obtain model parameters. The outcome variable of this regression is the fraction of trials with success (in this case, fraction of examinations with disease). Furthermore, increased confidence in illness rates, $y$, with increased number of health examinations can be accounted for with this regression. In both models, we assigned individuals to exposure categories to account for day-to-day variability and error or uncertainty in estimates of mean exposure and to capture the changing slope of the exposure-response relation.

### Results

Table 1 contains demographic information for individuals from the 55 study households, and mean (SD) number of health reports obtained. Figure 2 shows ARI (divided into AURI and ALRI)—mean fraction of weekly examinations in which an individual was diagnosed with either infection—stratified by demographic subgroup. Differences between female and male ARI rates by two-sided two-sample $t$-test were not statistically significant.

<table>
<thead>
<tr>
<th>Age group* (years)</th>
<th>Total (n=345)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5 (n=93)</td>
<td></td>
</tr>
<tr>
<td>5–14 (n=109)</td>
<td></td>
</tr>
<tr>
<td>15–49 (n=120)</td>
<td></td>
</tr>
<tr>
<td>≥50 (n=23)</td>
<td></td>
</tr>
<tr>
<td><strong>Girls/women</strong></td>
<td><strong>193</strong></td>
</tr>
<tr>
<td>Mean (SD) age at end of study</td>
<td></td>
</tr>
<tr>
<td>30.0 (1.4)</td>
<td></td>
</tr>
<tr>
<td>29.4 (10.0)</td>
<td></td>
</tr>
<tr>
<td>29.9 (9.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Mean (SD) health reports</strong></td>
<td></td>
</tr>
<tr>
<td>72.2 (23.9)</td>
<td></td>
</tr>
<tr>
<td>82.2 (16.3)</td>
<td></td>
</tr>
<tr>
<td>80.5 (17.7)</td>
<td></td>
</tr>
<tr>
<td>73.9 (19.1)</td>
<td></td>
</tr>
<tr>
<td>78.4 (19.7)</td>
<td></td>
</tr>
</tbody>
</table>

*We chose these age divisions because: children <5 years are most susceptible to acute respiratory infections; people ≥50 years begin to show chronic conditions; and people usually start work or marry at age 15 years. (From early 1997, to June, 1999.)

### Statistical analysis

Effect size of various factors for ARI and ALRI were estimated with two models. First, a linear model with an ordinary least-squares regression of illness rates. We accounted for clustering of observations in units of household, and used robust estimates of variance to correct for this and any statistical outliers in estimation of standard errors.

THE LANCET • Vol 358 • August 25, 2001

For personal use. Only reproduce with permission from The Lancet Publishing Group.
significant for children aged 4 years and under but were for other age groups: \( p = 0.02, 5-14 \) years; \( p = 0.001, 15-49 \) years; and \( p = 0.040, 50 \) or older. Differences between female and male ALRI rates were only significant in the 15-49 year subgroup (\( p = 0.001 \)). Significance was not sensitive to use of equal or unequal variances. The \( p \) values accounted for clustering of observations in units of household. Exposure to high emissions from cooking and other domestic activities for adults result in women being around twice as likely as men to be diagnosed with ARI or ALRI.

Figure 3 shows AURI and ALRI rates for infants and children, and young and adult individuals, against PM\(_{10}\) exposure category. We did not analyse the 50-years-and-older group because there were few people in the group and because we did not know their exposure history, which could have resulted in chronic diseases. Total ARI and ALRI rates increased at a higher rate for exposures to PM\(_{10}\) below 2000 \( \mu g/m^3 \) in both age groups.

Tables 2 and 3 show the effect estimates of various factors for ARI and ALRI estimated from the linear and logistic models, respectively. Results from both models confirm the relation shown in figure 3. Risk of ARI and ALRI increased with PM\(_{10}\) exposure but the rate of increase fell at exposures greater than 2000 \( \mu g/m^3 \). The concave shape of the exposure-response relation for ARI (and ALRI for 5-49 year age group) was confirmed by analysis with a continuous-exposure variable and inverse-quadratic relation (results not shown).

In the first 60 months after birth, consistent with previous work,\(^{23-26}\) the probability of being diagnosed with ARI and ALRI decreased by 0.009 (\( p = 0.08 \)) and 0.01 (\( p = 0.02 \)), respectively, with every year of age. People younger than 5 years were diagnosed in 8% more weeks (\( p = 0.001 \)) with AURI (5% for ALRI, \( p = 0.001 \)) than those aged 5-49 years, after controlling for exposure and other factors. After the age of 5 years, age increased the probability of being diagnosed with ALRI.

---

**Figure 3: Unadjusted exposure-response relation**

The shape of the relation was not sensitive to marginal modifications in exposure categories or to use of median, rather than mean, disease rates. See tables 2 and 3 for adjusted relation.

---

**Table 2: Adjusted excess risk of various factors for ARI and ALRI estimated with a linear model**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Age &lt;5 years</th>
<th></th>
<th>Age 5–49 years</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARI (n=93)</td>
<td>ALRI (n=93)</td>
<td>ARI (n=229)</td>
<td>ALRI (n=229)</td>
</tr>
<tr>
<td></td>
<td>Excess risk (95% CI)</td>
<td>( p )</td>
<td>Excess risk (95% CI)</td>
<td>( p )</td>
</tr>
<tr>
<td>PM(_{10}) exposure (( \mu g/m^3 ))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;200 (reference category)</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
</tr>
<tr>
<td>200-500</td>
<td>0.06 (0.002-0.10)</td>
<td>0.02*</td>
<td>0.01 (-0.01 to 0.03)</td>
<td>0.16*</td>
</tr>
<tr>
<td>500-1000</td>
<td>0.06 (0.002-0.11)</td>
<td>0.04*</td>
<td>0.01 (-0.01 to 0.03)</td>
<td>0.24*</td>
</tr>
<tr>
<td>1000-2000</td>
<td>0.13 (0.006-0.20)</td>
<td>0.01*</td>
<td>0.03 (0.00-0.06)</td>
<td>0.05*</td>
</tr>
<tr>
<td>2000-3500</td>
<td>0.14 (0.006-0.21)</td>
<td>0.01*</td>
<td>0.03 (-0.01 to 0.06)</td>
<td>0.16*</td>
</tr>
<tr>
<td>&gt;3500</td>
<td>0.18 (0.01-0.34)</td>
<td>0.04*</td>
<td>0.04 (-0.01 to 0.11)</td>
<td>0.30*</td>
</tr>
<tr>
<td>&gt;4000</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
</tr>
<tr>
<td>&gt;7000</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
</tr>
<tr>
<td>Female</td>
<td>-0.0007 (-0.04 to 0.04)</td>
<td>-0.009 (-0.03 to 0.01)</td>
<td>0.43</td>
<td>0.013 (-0.01 to 0.03)</td>
</tr>
<tr>
<td>Age†</td>
<td>-0.009 (-0.01 to 0.00)</td>
<td>0.08</td>
<td>-0.001 (-0.02 to 0.00)</td>
<td>0.002</td>
</tr>
<tr>
<td>Smoking</td>
<td>-0.00</td>
<td>-</td>
<td>-0.00</td>
<td>-</td>
</tr>
<tr>
<td>Village type‡</td>
<td>0.003 (-0.02 to 0.04)</td>
<td>0.02</td>
<td>0.006 (-0.02 to 0.04)</td>
<td>0.70</td>
</tr>
<tr>
<td>Number of people in household</td>
<td>0.0005 (-0.01 to 0.01)</td>
<td>0.01</td>
<td>0.0001 (-0.01 to 0.01)</td>
<td>0.99</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.20</td>
<td>-0.16</td>
<td>0.22</td>
<td>-0.17</td>
</tr>
</tbody>
</table>

\( \text{ARI}=\text{acute respiratory illness}, \text{ALRI}=\text{acute lower respiratory illness}. \text{The lowest exposure category (<200 } \mu g/m^3 \text{) was taken as the reference category. Therefore, the coefficients for all other exposure categories are fractional fraction-of-weeks with illness relative to this category. } * \text{Jointly significant (p}<0.03). \text{†Additional probability of diagnosis with illness with every additional year of age/extra person in house. } \text{‡Cattle compound or maintenance compound.} \)
possibly because of chronic effects of earlier exposure. In this age group, age did not affect the risk of ARI in the linear model, but in the logistic model a slight reduction of ARI with increasing age was seen which cannot be explained by known physiological mechanisms.

In independent analysis, smoking increased the risk of ARI by 0.02 (p=0.04) in the linear model and with an odds ratio of 1.48 (p=0.02, 95% CI 1.07–2.04) in the logistic model. The increase in ALRI risk from smoking was not significant because of small sample size. When analysed as a source of particulate matter and hence as contributing to exposure, the coefficient of smoking was no longer significant (p=0.47). Significance remained unchanged for all other variables. This result shows that effects of smoking on ARI might be similar to combustion products from other forms of biomass. At the same time, smoking has been causally linked with many other health hazards, some of which might be similar to other biomass products and others, in particular lung and other cancers, different.

Village type had no significant effect on rate of ARI after exposure and other factors were accounted for, which we attribute to similar incomes and diets. Furthermore, the type of village that a household resides in is decided by the ranch management and is not correlated with income. A portion of wages are paid as food rations (milk and staple food items), which further equalises food consumption among households. Number of people per household was not significant with respect to ARI, possibly because with the pastoralist life-style of the study group, activity patterns are a more important determinant of the amount of time occupants spent inside together than the number of household members. For the 5–49 age group, we accounted for the amount of household cooking tasks done (none, low, medium, high) and the intensity of exposure (concentration when a person was close to the stove and emissions were highest). The group who cooked most was diagnosed with ALRI in an additional 2% of weeks (p=0.03), but the remaining coefficients were not significant.

Exposure-response relation, when estimated only from average daily PM10 concentration and time spent indoors (ie, without adjustment for activities and movement patterns of individuals) showed that female individuals older than 5 years had an additional risk of ARI of 0.03 (p=0.001) and ALRI of 0.01 (p=0.003). This method of exposure estimation underestimates exposure of women, who cook more than men.24 In individuals younger than 5 years, sex was not significant (p=0.87 for ARI, p=0.47 for ALRI). A similar relation was seen in the logistic model (results not shown).

Controlling for the amount of cooking activity that a person did eliminated the significance of sex, confirming that the role of sex was a substitute for exposure patterns (ie, a proxy for the omitted variable of high-intensity exposure) when average daily PM10 concentration was used. This bias was further confirmed by noting that when estimating exposure with average daily PM10 concentration and time alone, the role of sex became important only after the age of 5 years when female family members actually take part in household activities. After inclusion of high-intensity exposure, sex was no longer an effective predictor of ARI or ALRI.

**Discussion**

We have shown that increased exposure to indoor PM10 increases the frequency of ARI. Rate of increase of exposure-response is highest for exposures below 1000–2000 μg/m3. Although this concave shape of the exposure-response relation fell inside 95% CI, it was confirmed by analysis with a continuous exposure variable. This result suggests that public-health programmes to reduce adverse impacts of indoor air pollution in developing countries should concentrate on measures that reduce average exposure to below 2000 μg/m3 as suggested by Bruce and colleagues.20

Health status of study individuals was likely to have been affected by treatment provided as part of the programme. Treatment was given for ethical reasons; it also standardised treatment in the study group, and prevented confounding from differing participant access to health-care facilities. However, if treatment had affected people differently in a way that was correlated with severity or exposure, the shape of the exposure-response curve would be modified. Thus, our analyses were done on the basis of the use of a small amount of health care.

Only a small fraction of variance was explained with our model. Birthweight, perinatal diseases, no breastfeeding, nutrition, and childcare practices are all additional risk factors for ARI.25 A life-history approach
that includes these factors could provide more information about ARI risk.

Exposure-assessment methods have commonly used average pollution levels. Indoor smoke exposure occurs in an episodic manner, hence use of average concentration biases calculation of exposure and health effects. Exposure intensity did not contribute to incidence of disease, once its role was accounted for in total exposure. However, for the highest exposure groups (notably the individuals who cook) around a half daily exposure occurs in high-intensity episodes. Thus reduction of total exposure by reduction of peak emissions remains important.

Technology-transfer programmes and public-health initiatives provide various benefits in less-developed nations. More than 2 billion people rely on biomass as their main source of energy, hence efforts to introduce new energy technologies should account for health outcomes. Many national, multilateral, and private donors have promoted improved (high-efficiency and low-emission) stoves. Many of these programmes, although lowering average emissions, might not have reduced exposure below 2000 μg/m³, let alone a few hundred μg/m³, which would provide substantial health benefits. The concave exposure-response relation that we found suggests that although improved wood stoves substantially reduce exposure, they give fewer health benefits than charcoal, which can reduce exposure to very low levels. Other energy technology changes from wood to charcoal, or to kerosene, gas, or electricity, also need to be evaluated for public health and environmental effects (e.g., effects on vegetation and greenhouse gas emissions). With a better quantitative understanding of health effects of indoor smoke, development, public health, and energy research can be aimed at reduction of acute respiratory infections.

Contributors
Majid Ezzati led the field research project, devised methods for exposure assessment and health data collection, analysed data, and wrote the report. Daniel Kammen designed the project, helped with field research design, data analysis, and writing the report.

Acknowledgments
S Munyi, J Murithi, the administration of Nanyuki District Hospital, and A W Murithi from Kenyatta National Hospital and National ARI Programme helped design and execute health monitoring. We thank Bernard Mbenda, Mark Egelian, Peter Ekuam, Mary Lokeny, and Jackson Ngisirkale for obtaining data; and residents of Mpala Ranch for their hospitality. African Academy of Sciences provided institutional support in Kenya. N Goldman, B H Singer, K R Smith, and two anonymous reviewers provided valuable comments on data-analysis methodology and epidemiology of ARI. Our research was supported by grants from the Summit and Compton Foundations, Social Science Research Council (SSRC) and Princeton University’s Council on Regional Studies and Center of International Studies (through a grant from the MacArthur Foundation).

References