PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The Peru Urban versus Rural Asthma (PURA) Study: methods and baseline quality control data from a cross sectional investigation into the prevalence, severity, genetics, immunology, and environmental factors affecting adolescent asthma in Peru
AUTHORS	Colin L Robinson, Lauren M Baumann, Robert H Gilman, Karina Romero, Juan M Combe, Lilia Cabrera, Nadia N Hansel, Kathleen Barnes, Guillermo Gonzalvez, Robert A Wise, Patrick N Breysse and William Checkley

VERSION 1 - REVIEW

REVIEWER	Philip W Ind Consultant Physician Hon Senior Lecturer NHLI IMperial College Hammersmith Hospital London UK No conflict of interest
REVIEW RETURNED	05/11/2011

THE STUDY	This is an original, well conducted and ambitious cross-sectional study examining asthma prevalence and severity in 2 communities, one urban (Lima) and one, hotter and wetter, rural region (Tumbes), in Peru. The primary outcome was self-reported asthma symptoms included in comprehensive questionnaires in 13-15 year olds. Multiple other data were collected but are not reported. Current asthma symptoms were reported in 12% of adolescents in Lima but in only 3% in Tumbes. Severe persistent symptoms were found in 5% of Lima asthmatics but in 14% of asthmatics in Tumbes although this was only in 3 out of 22 subjects. There is no discussion of the statistics. It is not clear how the figure of 1851 is arrived at in the Abstract ie what reflects a shortfall in random census sample in Lima and what reflects a shortfall in all invited Tumbes adolescents? There appears to be a conflict with the statement regarding sample size l. 203. The writing is not of a high standard even allowing for Americanisations of spelling, syntax, prepositions etc. There are multiple missing words eg l. 74 of Abstract, l. 105, l.149, l. 215, l. 366, l. 372, l.404, l.407, l.408, l.439, l.459, l.512, l.515, l.516, l.519, l.676. Multiple other errors eg 'understand or performing' l. 173 p8, 'bronchodilators' repeatedly, 'from' rather than 'for' l.224 'on' rather than 'in' eg l.291 also reverse ll. 389, 392, 393.
RESULTS & CONCLUSIONS	The manuscript really reflects work in progress as the vast majority of results are not reported. Much of the missing data will be of interest. As written, this paper deals with methodological considerations (eg spirometry performance, skin testing) and would

	be of practical interest to respiratory epidemiologists rather than a general audience. The key is missing to Figure 4 in the legend which should refer to panels rather than 'rows'. Detailed data from the asthma questionnaires in the 2 groups, relating to exposures, etc would be of interest. Some discussion of the validity of the census data is required. I am not sure that these experiences in conducting such a field tsudy can be generalised to other countries (l.532). Regarding the authors conclusions; no direct information on airway inflammation has been collected but rather on a surrogate (l.545) and evidence regarding pollution as a contribution can only be by correlation (l. 548).
GENERAL COMMENTS	Has the potential to be an important study when complete data are available.

REVIEWER	Luz Carbajal Arroyo Head Of Department Of Statistics, Demography, Humanities And Social Sciences Universidad Peruana Cayetano Heredia
REVIEW RETURNED	14/11/2011

THE STUDY	The statistical methods used are not described in methods
GENERAL COMMENTS	Compare the results found in the two areas and evaluate the differences.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:
Reviewer: Philip W Ind
Consultant Physician
Hon Senior Lecturer
NHLI Imperial College
Hammersmith Hospital
London UK

This is an original, well conducted and ambitious cross-sectional study examining asthma prevalence and severity in 2 communities, one urban (Lima) and one, hotter and wetter, rural region (Tumbes), in Peru. The primary outcome was self-reported asthma symptoms included in comprehensive questionnaires in 13-15 year olds. Multiple other data were collected but are not reported. Current asthma symptoms were reported in 12% of adolescents in Lima but in only 3% in Tumbes. Severe persistent symptoms were found in 5% of Lima asthmatics but in 14% of asthmatics in Tumbes although this was only in 3 out of 22 subjects.

R1Q1. There is no discussion of the statistics. The statistical methods used are not described in methods.
R1A1. This point is well-taken, however, this manuscript is meant to be a Protocol Paper in which we provide a comprehensive description of the study design, implementation, standard operating procedures, and quality control of primary outcome measures. In this manuscript, we only present simple means and percentages. To describe our specific aims for this study, we have added a description of the statistical methods under the Methods subheading “Specific Aims and statistical considerations” that describes in the five main aims of our study. However, statistical methods for each of the specific aims are presented in detail in other previous publications (Robinson CR et al Thorax 2011;66:1051-7 and Baumann LM et al. J Allergy Clin Immunol 2011;127:875-82) or will be

presented in subsequent publications. We have recently completed the analysis and final draft of the relation of total serum IgE to asthma and lung function, which we plan to submit for publication soon and are in the midst of conducting the laboratory analysis of vitamin D levels and candidate genes. These publications will be forthcoming.

R1Q2. t is not clear how the figure of 1851 is arrived at in the Abstract ie what reflects a shortfall in random census sample in Lima and what reflects a shortfall in all invited Tumbes adolescents?
R1A2. This is a good point. We have now noted how we arrived at the numbers of 725 in Lima and 716 in Tumbes given the starting figure of 1851. The first paragraph under the Results section now instead reads:
Of the 1851 potential participants listed on our censuses and approached across both sites, a total of 1441 adolescents were recruited to participate in the study, 725 in Lima and 716 in Tumbes. In Lima, 321 (30.4%) refused participation or were unable to be contacted, while 10 (0.9%) were ineligible, generally because their ages fell out of our study age range. In Tumbes, 67 (8.4%) refused participation or were unable to be contacted, while 12 (1.5%) were ineligible for similar reasons. Of those not enrolled from census, there were no significant differences in distribution by age or sex (data not reported).
As an additional note for the reviewers, the population in Lima was more difficult to enroll than that of Tumbes because in urban Lima there were more instances in which either the parent/guardian or child (or both) was not present in the home on multiple visits by study field workers. We further elaborate on this topic in fourth paragraph of the Discussion section.

R1Q3. There appears to be a conflict with the statement regarding sample size.
R1A3. This is a good point. We did not mean to give the impression that we had set a goal of exactly 720 participants at each site, but rather that we intended to enroll a number that fell within a very narrow window of 720. We have corrected this sentence to read, “Our initial assumption for loss to follow-up between visits was approximately 5%. Therefore, we aimed to enroll approximately 720 participants at each site, which we achieved as noted below.”

R1Q4. The writing is not of a high standard even allowing for Americanisations of spelling, syntax, prepositions etc. There are multiple missing words eg l. 74 of Abstract, l. 105, l.149, l. 215, l. 366, l. 372, l.404, l.407, l.408, l.439, l.459, l.512, l.515, l.516, l.519, l.676. Multiple other errors eg 'understand or performing' l. 173 p8, 'bronchodilators' repeatedly, 'from' rather than 'for' l.224 'on' rather than 'in' eg l.291 also reverse ll. 389, 392, 393.
R1A4. Thank you for taking such a detailed look at the spelling/grammar of our manuscript. We have made all the corrections you suggested, and upon your advice have further reviewed the manuscript for similar errors, which we also corrected. The only clarification of which we are unsure is the reference to repeated use of “bronchodilators,” as this is standard parlance in Pulmonary Medicine. We have changed this to the singular “bronchodilator” where appropriate, but we would appreciate further suggestion from the reviewer if this is not the edit desired.

R1Q5. The key is missing to Figure 4 in the legend which should refer to panels rather than 'rows'.
R1A5. We have added the appropriate information to the Figure description and changed the text so that it refers to panels rather than rows: “[F]rom all individuals in both sites (top two panels), all included in census (middle two panels), and all recruited (bottom two panels).”

R1Q6. Some discussion of the validity of the census data is required.
R1A6. Census data at our Lima site was originally obtained in 1997, as this neighborhood has for years served as a research site for us and our colleagues. The data was re-collected in 2000, and frequent updates have been made throughout the years. Each household was approached in door-to-door fashion, and census data on all permanent members of the household were obtained. Key data obtained included: full name, date of birth, and sex. From this data we performed our sample

selection. At our Tumbes site, there was a census that had recently been completed 4 months prior to the beginning of our data collection. The Tumbes site is a newly-created research site, with the initial project being the Proyecto de Eliminación de Cisticercosis (Cysticercosis Elimination Project). A map of each of the 23 villages in the Tumbes site was created, and as above, each permanent household member's data was recorded: full name, date of birth, and sex. This information is summarized under "Study Population" in the Methods section of the manuscript.

R1Q7. I am not sure that these experiences in conducting such a field study can be generalised to other countries (l.532).
R1A7. This is a good point. While our experience implies that in many resource-poor settings this quality and extent of data is possible to collect, this does not mean that in all resource-poor settings this can be accomplished. It does require a certain minimum level of infrastructure that allows for sufficient logistical support. Therefore, we have changed this portion of the text to read, "Nonetheless, our study shows that we can collect extensive high-quality data by home visitation in a resource-poor setting, and therefore this type and quality of data may be collected in other, though not necessarily all, resource-poor settings as well."

R1Q8. Regarding the authors conclusions; no direct information on airway inflammation has been collected but rather on a surrogate (l.545) and evidence regarding pollution as a contribution can only be by correlation (l. 548).
R1A8. Thank you for clarifying this point. Exhaled NO is a well-known marker of airway inflammation, as reported in many studies. Therefore, we have changed this sentence to read, "[M]easures of airflow obstruction and markers of airway inflammation can be successfully collected…" On your latter point regarding pollution as a contributor, we agree and have eliminated this clause from the conclusion paragraph.

R1Q9. Detailed data from the asthma questionnaires in the 2 groups, relating to exposures, etc would be of interest…The manuscript really reflects work in progress as the vast majority of results are not reported. Much of the missing data will be of interest. As written, this paper deals with methodological considerations (eg spirometry performance, skin testing) and would be of practical interest to respiratory epidemiologists rather than a general audience…Has the potential to be an important study when complete data are available.
R1A9. We have combined several related comments into one final comment here for the sake of ease of explanation. First, we appreciate the interest in the data collected by our research. To clarify for our reviewers, we have published two manuscripts (listed below) based on the data collected by the methods described in this current manuscript. The goal of this manuscript is to disseminate widely the extensive methods we utilized in our data collection and to report that this quality/quantity of data can be obtained in a resource-poor setting, which has not been done to this extent in previous studies. For readers of this manuscript, we have clarified this point in both the Methods and Discussion section.
Robinson CR, Baumann LM, et al. Effect of urbanisation on asthma, allergy and airways inflammation in a developing country setting. Thorax. Dec 2011; 66(12):1051-7.
Baumann LM, Robinson CR, et al. Effects of distance from a heavily transited avenue on asthma and atopy in a periurban shantytown in Lima, Peru. J Allergy Clin Immunol. Apr 2011;127(4):875-82.

Reviewer #2
Reviewer: LUZ CARBAJAL ARROYO
HEAD OF DEPARTMENT OF STATISTICS, DEMOGRAPHY, HUMANITIES AND SOCIAL SCIENCES, UNIVERSIDAD PERUANA CAYETANO HEREDIA

R2Q1. The statistical methods used are not described in methods. COMPARE THE RESULTS

FOUND IN THE TWO AREAS AND EVALUATE THE DIFFERENCES.

R2A1. We thank the reviewer for these comments. We would like to point out, however, that this manuscript is meant as a Protocol Paper, where we only summarize the study design, implementation and standard operating procedures, and provide quality control data for important outcome and exposure variables. Analyses of primary data and statistical methods used in these analyses have been or will be presented in detail elsewhere in other publications. For reasons of space, especially given the length of the current manuscript, we decided it would be better not to include the methods used in previous papers in this manuscript. For the purposes of this paper, we only used simple means and percentages. Under the Methods subheading "Specific Aims and statistical considerations", we know describe the five main aims of our study, and the statistical methods used in this Protocol Paper.

We would kindly ask this reviewer to refer to the following publications for a detailed description of statistical methods used in the primary analysis of data:

Robinson CR, Baumann LM, et al. Effect of urbanisation on asthma, allergy and airways inflammation in a developing country setting. Thorax. Dec 2011; 66(12):1051-7.

Baumann LM, Robinson CR, et al. Effects of distance from a heavily transited avenue on asthma and atopy in a periurban shantytown in Lima, Peru. J Allergy Clin Immunol. Apr 2011;127(4):875-82.

Specifically, in Robinson et al. published in Thorax earlier this year, we compare the prevalenc and risk factors of asthma between our populations in detail. We are including copies of these papers as supplementary material for use by the reviewers.

Several other publications of primary data analysis are forthcoming. We have recently completed the analysis and final draft of the relation of total serum IgE to asthma and lung function, which we plan to submit for publication soon and are in the midst of conducting the laboratory analysis of vitamin D levels and candidate genes.

VERSION 2 – REVIEW

REVIEWER	Philip W Ind Consultant Physician Hon Senior Lecturer NHLI IMperial College Hammersmith Hospital London UK No conflict of interest
REVIEW RETURNED	20/12/2011

THE STUDY	The important research questions which the PURA study aims to address will hopefully be answered once the genetic and other laboratory data are available. The current manuscript is better written with the aims of the paper summarised on p.8 but the study aims are well summarised on p.21 and the data for these are not yet available. The quality control data included are of little general interest and to a considerable extent implict in the published papers.
RESULTS & CONCLUSIONS	I congratulate the authors on the recent publication of 2 high quality useful publications related to the detailed questionnaire, spirometry, exhaled NO, skin test and indoor particulate data they have collected. The data in Lima were analysed to demonstrate an increased risk of

	asthma symptoms and atopy in relation to distance from a heavily transited avenue in the peri-urban shantytown.
REPORTING & ETHICS	I would not suggest that this represents redundant publication but there is no new data and all the information contained would be better included when the genetic and other results are available.