PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Determinants of self-reported smoking and misclassification during pregnancy, and analysis of optimal cut-off points of urinary cotinine: a cross sectional study
AUTHORS	Aurrekoetxea, Juan; Murcia-Hinarejos, Mario; Rebagliato, Marisa; Lopez, Maria; Castilla, Ane; Santa Marina, Loreto; Guxens, M; Fernández-Somoano, Ana; Espada, Mercedes; Lertxundi, Aitana; Tardón, Adonina; Ballester, Ferran

VERSION 1 - REVIEW

REVIEWER	Van Tong Epidmiologist Centers for Disease Control and Pregnancy USA I declare no competing interests
REVIEW RETURNED	09-Oct-2012

THE STUDY	Minor comments on the abstract that would help to clarify key findings of the study. See reviewer's comments.
GENERAL COMMENTS	The study objectives were 1) to assess the prevalence of self-reported smoking and the UC levels in a cohort of pregnant women; 2) to assess the prevalence of misclassification of maternal smoking status according to the most widely accepted cut-off point in the literature of 50 ng/ml, and to study maternal factors associated with both self-reported and misclassification of maternal smoking; and 3) to identify the optimal cut-off point for UC that best distinguishes smokers from non-smokers in our study sample, according to frequency of smoking (occasional or daily smokers) and SHS exposure. Given that metabolism of nicotine is altered during pregnancy, there is very little information on the optimal cut-points for cotinine to assess true smoking status particularly in the presence of high secondhand smoke (SHS) exposure for pregnant women. Most studies that use biochemical verification to assess true smoking status utilize cut-points determined for non-pregnant populations. Thus this study does contribute to this literature, and a strength of this study is determining optimal cut-points for pregnant smokers by level of smoking and varying levels of SHS exposure. Additionally the study include a relatively large sample of pregnant women (n=2290) in Spain. Major comments: • Year of pregnancy or birth was not incorporated into the adjusted analysis. The recruitment spanned 2003 to 2008. It would be important to include year of pregnancy in the model. If adding it does not make a difference in the adjusted ORs, then you can leave the

	tables as is but note that it was tested in the methods or discussion. • The cut-off for occasional smokers of 27ng/ml had a very poor positive predictive value (.164) (Table 4). I suspect this might be due to the high SHS exposure. It might be useful to examine occasional smoking by SHS exposure (thus exclude daily smokers). With such a low PPV, I would be weary to recommend urine cotinine as a valid biomarker for occasional smokers in populations where there is high SHS exposure. Some additional comments to consider: Abstract: • Objective: consider adding" best distinguishes daily and occasional smoking from non-smokers with varying levels of SHS exposure" Daily and occasional smoking is not mentioned in the abstract until the results. • Design: o "we used logistic regression models to study" o add the cut-point used for discriminate smoking status for UC (50ng/ML) o consider clarifying "The cut-points were also calculated after stratification among nonsmokers for SHS by number of sources." • The first sentence of conclusion is not clear. "Current efforts made to prevent smoking in pregnant women…." Makes it appear that you evaluated efforts in the study which you do not. Consider changing it to "Prevalence of smoking during pregnancy in Spain remains high." The second sentence , "UC is a reliable biomarker…" also does not follow from the study since it was not compared to another biomarker. Manuscript • P7, line 36, can you provide information on the time of day the urine samples were taken. Also I do not see any analysis looking at time of last cigarette and SHS to see whether this may explain discordance. • p8, line 12, clarify what type of distribution the Mann-Whitney U and Kruskal Wallis tests for…is it to test a linear trend? • P8, line 40, provide the n for the women who did not smoke but had UC levels >200ng/ml • P8, line 44, provide the n for the self-reported smoker and claims to stop smoking. Additionally provide n for the sample used to assess optimal cut-points. • P11, line 10, unclear what type of trend was observed. • P19, line 28, do you mean false negatives were due to maternal misreporting • P20, line 16, should confidence interval be 42 to 136 (not 50 to 100)? • P21, line 54, regarding the 27 ng/ml cut-point, please see previous comment.

REVIEWER	Roy M Nilsen, PhD, Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway
REVIEW RETURNED	26-Oct-2012

THE STUDY	There were no checklists attached to this manuscript.
GENERAL COMMENTS	The main objectives of the study by Aurrekoetxea et al were to estimate the prevalence of smoking among Spanish pregnant women, examine potential predictors for prenatal smoking, examine potential predictors for misclassification of self-reported smoking

	status, validate self-reported smoking status against urine cotinine cut-off levels of 50 ng/ml and 100 ng/ml, and estimate an optimal urinary cotinine cut-off from the current sample. The study sample used was the INMA cohort, a prospective multicenter study of Spanish pregnant women. The paper may add new insight to the field of misclassification of self-reported smoking in connection with pregnancy, but the paper needs considerable language editing before considering it for publication. Some additional comments are made below: 1. It is not clear from the manuscript at what gestational age the cohort subjects were excluded from the analyses. See for example, page 8, lines 40-51. A flow-chart showing exclusions from the initial sample size (n = 2644), accompanied with a description for each exclusion step, would help clarifying this. 2. Although analyses seem appropriate for the present study, the statistical analyses should be better described. The authors report that cut-offs were estimated and validated by using Youden's Index and ROC curves (page 8, lines 24-33). It would be useful to include some references and add a more precise description on how this was done. Others might want to use these methods to reproduce findings in similar populations. 3. In the statistics section, there is no mention of the level of significance (i.e., alpha) or the width of confidence intervals. It is not clear how the P value for linear trend was calculated for categorical variables. A reference and a more precise description of the bootstrap method should be included. Also, the authors should state whether the statistical tests were one or two-sided. 4. Table 1: It is not clear why the percentages in the two variables "Exposed to SHS…" and "Number of sources …." are not consistent. It would have been useful with an explanation for this in a footnote. Table 3: It is not clear why a number of prevalence estimates and odds ratios are omitted from the table for both adjusted and unadjusted analyses. All numbers should be included even though they are not statistically significant. 5. One of the paper's goals is to validate self-reported smoking status against various cut-offs of urinary cotinine levels. I am not sure why the authors excluded 53 self-reported nonsmokers when urinary cotinine levels were above 200 ng/ml. In my opinion, this subgroup of women is truly misclassified according to cotinine levels and should have been included in the estimation of the cut-off as well as in the analyses of sensitivity and specificity in Table 4. 6. Calculation of sensitivity and specificity for self-reported smoking were performed before and after exclusions of 277 women who had reported they had stopped smoking during pregnancy. I think this was done in order to examine the credibility of reporting among this subgroup, but this is not clear from the manuscript. At what gestational age did the women quit smoking (or reported quit smoking)? This information should be specified in the manuscript. 7. This is a methodological paper and several epidemiologic terms are used throughout the manuscript without providing explanations or definitions (i.e., internal validity, sensitivity, specificity, AUC etc.). I suggest that the authors include some definitions to improve

	readability. 8. The authors might want to discuss pros and cons using urinary cotinine levels instead of plasma/serum cotinine levels in validation studies. 9. Abstract: The abbreviation INMA is not defined. The study period should be indicated.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Van Tong
Epidmiologist
Centers for Disease Control and Pregnancy
USA
I declare no competing interests

The study objectives were 1) to assess the prevalence of self-reported smoking and the UC levels in a cohort of pregnant women; 2) to assess the prevalence of misclassification of maternal smoking status according to the most widely accepted cut-off point in the literature of 50 ng/ml, and to study maternal factors associated with both self-reported and misclassification of maternal smoking; and 3) to identify the optimal cut-off point for UC that best distinguishes smokers from non-smokers in our study sample, according to frequency of smoking (occasional or daily smokers) and SHS exposure.

Given that metabolism of nicotine is altered during pregnancy, there is very little information on the optimal cut-points for cotinine to assess true smoking status particularly in the presence of high secondhand smoke (SHS) exposure for pregnant women. Most studies that use biochemical verification to assess true smoking status utilize cut-points determined for non-pregnant populations. Thus this study does contribute to this literature, and a strength of this study is determining optimal cut-points for pregnant smokers by level of smoking and varying levels of SHS exposure. Additionally the study include a relatively large sample of pregnant women (n=2290) in Spain.

Major comments:
• Year of pregnancy or birth was not incorporated into the adjusted analysis. The recruitment spanned 2003 to 2008. It would be important to include year of pregnancy in the model. If adding it does not make a difference in the adjusted ORs, then you can leave the tables as is but note that it was tested in the methods or discussion.
- According to reviewer's suggestion, we have included the year of urine sampling in tables 1 and 3. Although univariate analysis showed a strong association between year of sampling and self-reported smoking, this association was no longer significant after adjusting for covariates.

• The cut-off for occasional smokers of 27ng/ml had a very poor positive predictive value (.164) (Table 4). I suspect this might be due to the high SHS exposure. It might be useful to examine occasional smoking by SHS exposure (thus exclude daily smokers). With such a low PPV, I would be weary to recommend urine cotinine as a valid biomarker for occasional smokers in populations where there is high SHS exposure.
- The prevalence of occasional smoking in our sample was 0.021 (37/1792) [pre-test probability]. Post-test probability, PPV, was higher, 0.162. We add in the text that excluding SHS exposed among non smokers this PPV improved to 0.412, with a cut-off point of 19ng/ml.

Some additional comments to consider:
Abstract:
• Objective: consider adding” best distinguishes daily and occasional smoking from non-smokers with

varying levels of SHS exposure" Daily and occasional smoking is not mentioned in the abstract until the results.
- This change has been made.
• Design:
o "we used logistic regression models to study"
- This typo has been corrected.
o add the cut-point used for discriminate smoking status for UC (50ng/ML)
- This change has been made.
o consider clarifying "The cut-points were also calculated after stratification among nonsmokers for SHS by number of sources."
- We add "by number of sources".
• The first sentence of conclusion is not clear. "Current efforts made to prevent smoking in pregnant women…." Makes it appear that you evaluated efforts in the study which you do not. Consider changing it to "Prevalence of smoking during pregnancy in Spain remains high." The second sentence , "UC is a reliable biomarker…" also does not follow from the study since it was not compared to another biomarker.
- We leave the first sentence of the conclusion. We add "Prevalence of smoking during pregnancy in Spain remains high". In the second sentence we change tobacco consumption by smoking status. We consider that it does not mean that we are comparing with other biomarkers.

Manuscript
• P7, line 36, can you provide information on the time of day the urine samples were taken. Also I do not see any analysis looking at time of last cigarette and SHS to see whether this may explain discordance.
- We add that samples were collected "in the morning". We add also that "No information about last cigarette or last SHS exposure was obtained".
• p8, line 12, clarify what type of distribution the Mann-Whitney U and Kruskal Wallis tests for…is it to test a linear trend?
- This sentence has been clarified in the text.
• P8, line 40, provide the n for the women who did not smoke but had UC levels >200ng/ml
- N added for the women who did not smoke but had UC levels >500ng/ml
• P8, line 44, provide the n for the self-reported smoker and claims to stop smoking. Additionally provide n for the sample used to assess optimal cut-points.
- N added for the self-reported smoker and claims to stop smoking
• P11, line 10, unclear what type of trend was observed.
- We changed the sentence in the text
• P19, line 28, do you mean false negatives were due to maternal misreporting
- In our study self reported smoking is treated as the reference, necessary for search the cut-off point. UC levels above de cut-off point and non reporters are, for us, false positives.
• P20, line 16, should confidence interval be 42 to 136 (not 50 to 100)?
- This typo has been corrected.
• P21, line 54, regarding the 27 ng/ml cut-point, please see previous comment.
- We add the 95% CI (11 to 43)

Reviewer: Roy Nilsen, biostatistician, PhD, Department of Public Health and Primary Health Care, University of Bergen, Norway.
Liv Kvalvik, MD, PhD research fellow, Department of Public Health and Primary Health Care, University of Bergen, Norway.
We have no competing interests.

The main objectives of the study by Aurrekoetxea et al were to estimate the prevalence of smoking among Spanish pregnant women, examine potential predictors for prenatal smoking, examine

potential predictors for misclassification of self-reported smoking status, validate self-reported smoking status against urine cotinine cut-off levels of 50 ng/ml and 100 ng/ml, and estimate an optimal urinary cotinine cut-off from the current sample. The study sample used was the INMA cohort, a prospective multicenter study of Spanish pregnant women. The paper may add new insight to the field of misclassification of self-reported smoking in connection with pregnancy, but the paper needs considerable language editing before considering it for publication. Some additional comments are made below:

- As we said in the cover letter, if you consider that we must refer the paper to a native English speaker to correct, please, let us know so that we could make corrections.

1. It is not clear from the manuscript at what gestational age the cohort subjects were excluded from the analyses. See for example, page 8, lines 40-51. A flow-chart showing exclusions from the initial sample size (n = 2644), accompanied with a description for each exclusion step, would help clarifying this.

- Following reviewer's suggestion, a flow-chart showing the exclusions have been done.

2. Although analyses seem appropriate for the present study, the statistical analyses should be better described. The authors report that cut-offs were estimated and validated by using Youden's Index and ROC curves (page 8, lines 24-33). It would be useful to include some references and add a more precise description on how this was done. Others might want to use these methods to reproduce findings in similar populations.

- We have included a more detailed explanation of the methods used in the analysis of the optimal cut-off point for UC. Some references have also been added.

3. In the statistics section, there is no mention of the level of significance (i.e., alpha) or the width of confidence intervals. It is not clear how the P value for linear trend was calculated for categorical variables. A reference and a more precise description of the bootstrap method should be included. Also, the authors should state whether the statistical tests were one or two-sided.

- The level of confidence (1-α=95%) used in the confidence intervals has been added in the statistics section. All the tests performed in the manuscript were two sided, this has been also specified in the text.

- As we said above we changed in the text "a clear trend" by "a clear dose-response pattern (not statistically tested)".

- The specific bootstrap method used, the R package, and a proper reference has been also added in the statistics section.

4. Table 1: It is not clear why the percentages in the two variables "Exposed to SHS…" and "Number of sources …." are not consistent. It would have been useful with an explanation for this in a footnote. Table 3: It is not clear why a number of prevalence estimates and odds ratios are omitted from the table for both adjusted and unadjusted analyses. All numbers should be included even though they are not statistically significant.

- We include a footnote in Table 1 "the percentages calculated including non exposed women".

- In Table 3, the variables statistically associated with the dependent variable of any of the models we entered them to the other models. Thus, the readability improves. There are two variables; social class and year of sampling that are close associated with the level of education and cohort, respectively. The inclusion in the model biased the estimates. We propose to leave the table in that way or, alternatively, exclude both variables in the table and mention them in the text.

5. One of the paper's goals is to validate self-reported smoking status against various cut-offs of urinary cotinine levels. I am not sure why the authors excluded 53 self-reported nonsmokers when urinary cotinine levels were above 200 ng/ml. In my opinion, this subgroup of women is truly misclassified according to cotinine levels and should have been included in the estimation of the cut-off as well as in the analyses of sensitivity and specificity in Table 4.

- This commentary deals with the trickiest issue of the article. In order to obtain an optimal cut-off point for urinary cotinine to distinguish smokers from non-smokers we needed a gold standard to compare with. In the absence of an 'objective measure' (independent of cotinine to avoid bias) the gold standard used was the response from questionnaire. This approach has also been applied

explicitly in other studies (Higgins et al, 2007; Benowitz NL et al, 2009; Hegaard et al, 2007; Seccareccia et al, 2003). The limitations of this approach have been discussed in detail in the "Limitations" section of the manuscript, and are mainly due to the risk of underreporting of the smoking habit. With the aim of dealing with this issue, we excluded self-reported non-smokers with implausible high UC levels, defined as a UC above 200 ng/ml. A similar strategy has also been applied by Hegaard et al (2007).
- However, we cannot assure that some non smoking women highly exposed to SHS truly have a UC>200 ng/ml, since the upper limit of the confidence intervals for the optimal cut-off point estimated for these women are even higher (201 in the case of women exposed to SHS in 2 or more environments). If this occurs, their exclusion would lead to a shift towards the left in the distribution of UC among non smokers, and a slight underestimation of the optimal cut-off point and a slight overestimation of the specificity as a result.
- We have therefore recalculated the table 4 using a more conservative cut-off point of 500 ng/ml, which is clearly in the range of smoking women. With the updated results, we can state that, if some degree of misclassification persists, it would be related with non declared smoking and would lead to a shift towards the right in both distributions, resulting in a slight overestimation of the optimal cut-off point and a probable underestimation of the specificity and sensibility.
- Some details have been added in the 'Statistical analysis' section of the manuscript..
- References:
- Higgins ST, Heil SH, Badger GJ, et al. Biochemical verification of smoking status in pregnant and recently postpartum women. Exp Clin Psychopharmacol 2007; 15:58-66.
- Benowitz NL, Bernert JT, Caraballo RS, et al. Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. Am J Epidemiol 2009; 169:236-48.
- Hegaard HK, Kjaergaard H, Møller LF, et al. Determination of a saliva cotinine cut-off to distinguish pregnant smokers from pregnant non-smokers. Acta Obstet Gynecol Scand 2007;86:401-6.
- Seccareccia F, Zuccaro P, Pacifici R, et al. Serum cotinine as a marker of environmental tobacco smoke exposure in epidemiological studies: The experience of the MATISS project. Eur J Epidemiol 2003;18:487-92.
6. Calculation of sensitivity and specificity for self-reported smoking were performed before and after exclusions of 277 women who had reported they had stopped smoking during pregnancy. I think this was done in order to examine the credibility of reporting among this subgroup, but this is not clear from the manuscript. At what gestational age did the women quit smoking (or reported quit smoking)? This information should be specified in the manuscript.
- Most women quit smoking during the first trimester of pregnancy (61%). We have monthly information about smoking status; however we have omitted this information from tables 1 and 3 to improve their readability. Taking into account the half-life of cotinine in plasma, these cases should not be related with UC. This group appears to be more prone to misreporting of smoking status and that is the reason because they were excluded in sensitivity analysis.
7. This is a methodological paper and several epidemiologic terms are used throughout the manuscript without providing explanations or definitions (i.e., internal validity, sensitivity, specificity, AUC etc.). I suggest that the authors include some definitions to improve readability.
- We have included in the text some definitions of those epidemiological concepts to improve readability.
8. The authors might want to discuss pros and cons using urinary cotinine levels instead of plasma/serum cotinine levels in validation studies.
- We have included in the text based in the references above “There are not validation studies of cotinine in different biological matrices, blood (plasma or serum), saliva or urine, so it cannot be established which the most reliable biomarker is”.
- References:
- Gorber SC, Schofield-Hurwitz S, Hardt J, et al. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. Nicotine Tob

Res 2009;11:12-24.
- SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002;4:149–59.
9. Abstract: The abbreviation INMA is not defined. The study period should be indicated.
- This both sentences have been clarified in the text.

VERSION 2 – REVIEW

REVIEWER	Roy M Nilsen, PhD, Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway
REVIEW RETURNED	03-Dec-2012

THE STUDY	I did not find any checklist
GENERAL COMMENTS	The manuscript is improved and most of our previous comments have been taken into account. However, regarding sensitivity and specificity analyses, it is still unclear from the manuscript why the authors exclude 35 self-reported nonsmokers when urinary cotinine levels are above 500 ng/ml. In my opinion, this subgroup of women is truly misclassified according to cotinine levels and should have been included in the estimation of sensitivity and specificity in Table 4. In other words: the authors have validated self-reported smoking habits using a more homogenous material confined to less misclassification than seen in the original material. As such, both sensitivity and specificity for these cut-offs may be spuriously high.

VERSION 2 – AUTHOR RESPONSE

We have change the Table 4 including all pregnant women. We added in a supplementary table the sensitivity analysis doing three alternative exclusions. Thus, the reader could be search the easier information in the table 4, and he could search complementary information about the stability of the validation parameters.