Author's response to reviews

Title: Translation and validation of Berlin Questionnaire in primary health care in Greece.

Authors:

Izolde Bouloukaki Dr (izolthi@gmail.com)
Ioannis D Komninos Dr (yiankom@yahoo.co.uk)
Charalampos Mermigkis Dr (mermigh@gmail.com)
Katerina Micheli Mrs (katerinamicheli@yahoo.gr)
Maria Komninou Mrs (mariakomninou@gmail.com)
Violeta Moniaki Mrs (vmoniaki@yahoo.gr)
Eleni Mauroudi Mrs (elenima23@hotmail.com)
Nikolaos M Siafakas Prof (siafak@med.uoc.gr)
Sophia E Schiza Dr (schiza@med.uoc.gr)

Version: 2 Date: 28 November 2012

Author's response to reviews: see over
Dear Editor,

Attached is the revised manuscript entitled “TRANSLATION AND VALIDATION OF BERLIN QUESTIONNAIRE IN PRIMARY HEALTH CARE IN GREECE”. The manuscript has been revised according to Reviewers suggestions. We would like to thank them for their valuable comments and we hope that the revised manuscript is suitable for publication.

On behalf of all authors,
Sincerely,

Izolde Bouloukaki. MD, PhD
Sleep Disorders Unit,
Dept. of Thoracic Medicine,
University of Crete

Reviewer's 1 report:
The authors have looked at the utility of a Greek version of the BQ in diagnosis of OSA. Under diagnosis of OSA remains an important issue. The manuscript is well thought out and well written.

Thank you for your comment.

Major Compulsory Revisions
There needs to be some review of the conclusions in the abstract. I'm not sure I agree that this version of the BQ is a valid and reliable instrument for identifying those with OSA. As a screening questionnaire you seek a low false negative rate and at an AHI > 5 the sensitivity is only 76%. The original Netzer paper had a sensitivity of 86% Specificity 77%. Id would like to see discussion around how this BQ might be used and a little more measured commentary. At all 3 cut points the sensitivity and specificity has limitations, neither ruling in or out OSA with confidence. What about the 24% false negatives at an AHI > 5? Do they matter? Is that level of sensitivity adequate? Can the authors explain what the sensitivity and specificity does not change very much at AHI > 5, 15, 30. In the Netzer manuscript Specificity increased as the AHI threshold increased and sensitivity fell. Why has that not occurred here?

Thank you for your comment. We agree with you and we have changed the conclusion of the abstract. As already mentioned, in the discussion section of the manuscript, our investigation showed that, BQ demonstrated moderate to high sensitivity and low to moderate specificity. Nevertheless, apart from the Netzer study, previous studies exploring the BQ as an instrument to detect sleep apnoea showed similar or lower sensitivity and specificity compared to our results. However, positive predictive value is high (80-94%), indicating that if BQ is positive, for example at an AHI>5, there is a high (94%) likelihood that a person would actually have sleep apnea. That’s why we suggest the use of BQ as a screening tool. As sleep apnoea frequently goes undiagnosed, primary care practitioners require a practical and sensitive screening tool to identify patients at high risk of having OSAS.

It is true that the sensitivity and specificity does not change very much at AHI > 5, 15, 30, compared to the Netzer manuscript. This may be due to several reasons. First, the population could be different. Polysomnography was offered to both high risk and
lower risk patients, but, as mentioned in the limitation section, more patients at high risk, who had sleep symptoms, selectively consented to the overnight polysomnography. Secondly, in the Netzer study, portable monitoring was used to assess the validity of the risk grouping strategy, compared to attended overnight polysomnography, which is the gold standard for the diagnosis of OSAS. Lastly, we have already mentioned in the discussion section, that other studies have shown similar results.

Comments related to the above were added in the revised manuscript.

*Can the authors please refine the methods around how they scored sleep and specifically respiration. Given the AHI is the gold standard measurement it must be absolutely clear how it was scored as different scoring techniques change the AHI substantively and likely explain some of the differences above. The reference they quote for scoring respiratory events is wrong, R&K was not for respi events, (Ref 14) and furthermore whilst they say the used AASM 2007 criteria they don’t define if they used Alternative or recommended criteria. Indeed the scoring system they’ve used for hypopneas seems to combine Recommended and Alternate criteria. Furthermore these scoring systems are based on AASM consensus, not ATS (page 7 line line 5) describe. This needs tightening and revision*

Thank you for your comment. We included a single standardized hypopnea definition, according to the AASM standard criteria. A comment related to the above was added in the revised manuscript.

*Can the authors define how they recruited people in Primary Care. Were they induced for example but being told it was a snoring and OSA study, and does that explain the high OSA prevalence? I think the most likely explanation for the high prevalence is the relative sensitivity of the scoring system making AHI > 5 not really the right cut point for OSA diagnosis. Ruehland Sleep 2009 describes the impact of different scoring systems on AHI very well and should be quoted.*

We recruited consecutive primary care patients who visited the General Hospital/Health Centre of Neapolis, Crete, Greece who met the inclusion criteria and completed the Greek version of the BQ. As explained in the limitation section, there may have been self-selection by the patients because those who had sleep symptoms might have selectively consented to the overnight polysomnography (129 of the 189 patients), explaining the high prevalence of OSAS in our group (118 patients).

It is true that using different published standard hypopnea definitions leads to marked differences in AHI. The scoring system, as was mentioned in the previous comments, included a single standardized hypopnea definition, according to the AASM standard criteria, therefore the high prevalence could not be attributed to the relative sensitivity of the scoring system.

*Minor Revisions*

diurnal somnolence is not assessed by the ESS, rather excessive daytime sleepiness (top of page 6)

We corrected it.
Reviewer's 2 report:
The need to develop expeditious and cost-effective tools to screen Obstructive Sleep Apnea Syndrome (OSAS) remains a challenge and an issue worthy of on-going investigation. In this study the authors want to validate a Greek translation of the Berlin Questionnaire (BQ) for OSAS and to explore whether this screening questionnaire could be used to help identify primary care patients at greater risk of having OSAS. The study was performed on 189 OSA patients. The authors conclude that based on their data, collected on mostly obese subjects, is a valid and reliable instrument for identifying patients at risk for OSAS in primary health care in Greece. The findings confirm that such screening tools should be used by primary care clinicians for OSAS prediction.
The title of the manuscript was clear in regards to the intent of the study and the question well defined. The abstract has a clear background and conclusion. There are some concerns.

Major Compulsory Revisions
1. The methods are mostly appropriate, however, it must be better clarified which subjects underwent nocturnal polysomnography. It is not quite so simple for the community clinician to reproduce not well defined rules with particular characteristics of population. The population analyzed by Netzer et al. had a BMI <30 kg/m2 in 370 (49.7) and ≥30 kg/m2 in 276 (37.1) subjects. Your 189 subjects had a BMI of 35.0 ± 25.1 kg/m2 (103 were obese). It is relevant to better report in Table 1 characteristics of the population studied (normal, overweight, obese, ...). A wide range of BMI is what we expect in a primary care population, and is relevant for the influence of different pathophysiological mechanism of upper airway obstruction during sleep. The reasons for excluding some participants from the sleep study are unclear. For example, why were non-obese so few in the study? PSG study was performed in 129 out of 189 patients (68.3%) and the diagnosis of OSAS was confirmed in 91.5% of these (n = 118).

Thank you for your comment. We agree with you and we have changed the second paragraph in the Methods section. The group of patients that met inclusion criteria was referred to the Sleep Disorders Unit, for evaluation of suspected sleep-disordered breathing and underwent nocturnal polysomnography. Of the 189 subjects who met the inclusion criteria, 129 underwent polysomnography. There was no specific reason for excluding some participants from the sleep study. These 60 subjects did not consent to complete a sleep study for personal reasons (for example inability to transfer from home to sleep unit).

Of the 189 subjects, 30 (15.9%) were normal, 56 (29.6%) were overweight and 103 (54.5%) were obese. We included these characteristics in table 1, as suggested. Non-obese were fewer in the study, because overweight and obese might have selectively consented to participate in this study and undergo the overnight polysomnography. On the other hand non-obese, especially in rural areas, are resistant to seeking medical help, because they consider themselves as healthy.

Comments related to the above were added in the revised manuscript,
2. Only high risk subjects at the questionnaire underwent nocturnal PSG? What about low risk subjects? This part of the methods is a major limitation and must be reported in the Results and discussed in the Discussion.

We agree with you comment, and this sentence was changed in the methods section, as already mentioned in the previous comment. Both low and high risk subjects at the questionnaire underwent nocturnal PSG.

3. Predictive results are referred only to subjects with specific questionnaire results. Please clarify what happened to the remaining 71 subjects. The 118 included all 103 obese subjects? This must be clearly stated through the paper. Respiratory disturbances during sleep was offered to both high-risk and lower-risk patients in the paper by Netzer et al.

Predictive results were referred only to subjects who fulfilled the PSG study (129 of the 189 patients) and not in those with specific questionnaire results. The remaining 60 subjects denied PSG study and were not further analyzed. This is clarified in the manuscript. Respiratory disturbances during sleep were offered to both high-risk and lower-risk patients, as in the paper by Netzer et al.

**Minor Essential Revisions**

1. The 2nd paragraph needs to be more focused on the main topic of the paper. The topic is a screening and not a diagnostic tool. As the authors well know, there have been a multitude of techniques purposed to be "screening" tools for OSAS. However, overnight study is still necessary to diagnose sleep apnea.

The recommended change was made in the revision.

2. Please use Berlin questionnaire (BQ) just the first time and than BQ is enough.

We corrected it.

3. Page 6: "diurnal somnolence ...... (ESS)". Was this and high blood pressure measured used to evaluate patients risk or just to characterize the population?

The BQ addresses items for the presence and frequency of snoring behaviour, wake time sleepiness or fatigue, and history of obesity and hypertension. How ESS is reported in the text is confounding. It must be better distinguished between values that characterize the population studied from the use of BQ.

ESS and blood pressure were used just to characterize the population.

**Discretionary Revisions**

Tables need to be improved. Table 2: Is not clear. It must report the findings on the 129 patients.

We changed Table 2, as suggested.