Screening methods and diagnosis of obstructive sleep apnoea in severely obese pregnant women

Authors
Heather Longworth\textsuperscript{a}, Kate McCallin\textsuperscript{b}, Ram Prakash Narayanan\textsuperscript{c}, Mark A. Turner\textsuperscript{d}, Siobhan Quenby\textsuperscript{e}, David Rycroft\textsuperscript{d}, Margaret Charnley\textsuperscript{f}, Julie Abayomi\textsuperscript{f}, Joanne Topping\textsuperscript{a}, Andrew D. Weeks\textsuperscript{d}, John PH Wilding\textsuperscript{c}

\textsuperscript{a}Liverpool Women's NHS Foundation Trust, Liverpool, UK; \textsuperscript{b}School of Medical Education, University of Liverpool, Liverpool, UK; \textsuperscript{c}Department of Obesity and Endocrinology, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK; \textsuperscript{d}Department of Women's and Children's Health, Institute of Translational Medicine, University of Liverpool, Liverpool, UK; \textsuperscript{e}Lancaster University, Lancaster Medical School, Clinical Research Hub, Lancaster, Lancashire, UK; \textsuperscript{f}Faculty of Education, Health & Community, Liverpool John Moores University, Liverpool, UK;

Keywords: High–risk pregnancy, obesity, pregnancy, morbidity, preeclampsia

Running title: Diagnosing obstructive sleep apnoea in pregnancy

Corresponding author:
Professor John Wilding
Clinical Sciences Centre
University Hospital Aintree
Longmoor Lane
What is already known about this subject?

- Obstructive sleep apnoea in the severely obese pregnant population is under-researched yet has potentially severe maternal and fetal consequences.
- Data regarding prevalence of OSA in pregnancy is scarce, although the estimated rate in women of reproductive age is 5-6%.
- The reliability of screening methods is variable, particularly when used in pregnant women.

What this study adds?

- 12% of 162 obese pregnant women scored positively on the Epworth Sleepiness Scale (ESS), but this was not confirmed on formal testing using more invasive techniques.
- Neither the ESS questionnaire nor the RUSleeping (RUS) meter appeared to provide a reliable method with which to diagnose OSA in pregnancy.
- Larger studies using screening questionnaires coupled with updated and less invasive monitoring technology are required before prevalence of OSA in pregnancy can be reliably estimated.
Abstract

Introduction

Obstructive sleep apnoea is an often-overlooked diagnosis, more prevalent in the obese population. Screening method accuracy, uptake and hence diagnosis is variable. There is limited data available regarding the obese pregnant population, however many studies highlight potential likely risks of apnoeic episodes, both to the mother and fetus, including hypertension, diabetes and pre-eclampsia.

Materials and Methods

162 women with a BMI≥35 were recruited from a tertiary referral hospital in the North West of England. They were invited to attend three research antenatal clinics, completing an Epworth sleepiness scale (ESS) questionnaire at each visit. A monitor to measure the apnoea hypopnoea index (AHI) was offered at the second visit. Data taken from consent forms, hospital notes and hospital computer records were collated and anonymised prior to statistical analysis.

Results

12.1% of women had an ESS score of ≥10, suggesting possible OSA. Rates increased throughout pregnancy although unfortunately there was a high rate of attrition. 37.6% of women used the RUSleeping meter; only one (2.1%) met prespecified criteria for OSA (AHI≥15). This individual had OSA categorised as severe and underwent investigations for preeclampsia, eventually delivering by emergency caesarean section due to foetal distress.
Conclusions

The accuracy of the ESS questionnaire to detect OSA in the pregnant population remains unclear. Further research on a larger sample size using more user-friendly technology to measure AHI would be beneficial. There are currently no guidelines relating to screening for OSA in the obese pregnant population, yet risks to both mother and foetus are well researched.

Abbreviations

Obstructive sleep apnoea (OSA), Apnoea hypopnea index (AHI), Epworth sleepiness scale (ESS), Body Mass Index (BMI), American Academy of Sleep Medicine (AASM)

Key Message

Obstructive sleep apnoea in the severely obese pregnant population is under-researched yet has potentially severe maternal and fetal consequences. Simple screening methods leading to early diagnosis are likely to be beneficial.
Main text

Sleep disordered breathing (SDB) covers a spectrum of symptoms ranging from snoring and upper airway resistance syndrome to obstructive sleep apnoea (OSA). Diagnosis of OSA depends on the average number of apneas and hypopneas per hour of sleep, also known as the Apnea Hypopnea Index (AHI). OSA is diagnosed when there are clinical symptoms plus AHI ≥ 5. Although there have been many studies into OSA in the general population, there is limited information regarding the prevalence and usefulness of screening tools to detect OSA in pregnant women. (1,2)

Classical symptoms include excessive daytime sleepiness, gasping or choking sensations during sleep and witnessed apneas. Daytime symptoms such as drowsiness are rather non-specific. In pregnancy this is further complicated as many women experience reduced quality sleep and impaired daytime functioning for other physiological reasons i.e. nocturia, discomfort, pain and general restlessness. (1,3) As a consequence, many believe that poor sleep and excessive daytime somnolence is normal during pregnancy, which can delay or even prevent presentation to a health care professional and the subsequent investigation into these symptoms. (4) Compared with men, women are more likely to mention lack of energy or fatigue when describing symptoms, and this can exaggerate scores on screening questionnaires. Depression and hypothyroidism can also increase ESS scores, which further complicates diagnosis. (5,6)

There are a number of mechanisms believed to increase the risk of OSA in pregnancy. OSA has a strong association with increased BMI, as does progressive weight gain in pregnancy. As the fetus grows, it displaces the diaphragm upwards and reduces lung functional residual capacity by up to 20%. Furthermore increased oestrogen levels
may lead to a narrowing of the upper airway due to vasomotor rhinitis, nasopharyngeal mucosal oedema and hyperaemia. (1,7)

Data regarding prevalence of OSA in pregnancy is scarce, and accurate rates are not yet known. Prevalence in women of reproductive age is estimated to be between 5-6%. (1) A small, prospective cohort study of 105 pregnant women in America showed 10.5% had OSA in first trimester and 26.7% in third, using formal overnight polysomnography testing. Of the 28 women with definitive OSA in the third trimester, it was mild in 23, moderate in 4 and severe in 1. Women were not statistically representative of the general obstetric population but did have equally distributed BMI values. (8,9)

Although pregnancy is thought to increase risk of OSA, Sarberg et al studied 180 women using the Epworth Sleepiness Scale (ESS) and ‘whole night respiratory recordings’ and demonstrated lower rates of OSA in pregnant (1%) versus a control group (3%) of age and BMI-matched non-pregnant women. Interestingly, the ESS median score was 7 in non-pregnant group and 9 in pregnant group (p<0.001) suggesting that the ESS may overestimate the rate of OSA when used for the screening of the pregnant population. It is important to recognize however, that the majority of women in this study were non-obese. (10)

OSA in pregnancy carries significant risks maternal and fetal risks, including intrauterine growth restriction, pulmonary hypertension, gestational diabetes mellitus, and preeclampsia. (11-15) For example, Chen et al analysed data from 791 women with OSA and 3955 randomly selected women without OSA. They found that
pregnant women with OSA were at increased risk for having preeclampsia, as well as growth restricted and preterm babies. (11)

There are a number of screening questionnaires available for detecting those at risk of OSA including the ESS, Berlin, Stop and Stop-Bang questionnaires. Although questionnaires are both economical and accessible, they have had variable success in previous studies. A comparative study of all four using 234 patients referred to a sleep clinic demonstrated high sensitivity rates for the Berlin, Stop and Stop-Bang questionnaires (95%, 91% and 98%) but low specificity rates (25% 25% and 26%). The ESS had the highest specificity (75%) but lowest sensitivity (73%). (16) Other studies have shown better performance of questionnaires, particularly in the second and third trimesters. (17)

ESS scores tend to increase throughout pregnancy. Sarberg et al found the mean ESS score increased from 7.9 to 8.7 from 1st to 3rd trimester (p<0.001.) (18) Similar outcomes were found by Hutchison et al: 4% had ESS score ≥10 pre-pregnancy, compared to 33% in final week of pregnancy. (4) Pien et al found the ESS scores increased significantly at all subsequent assessments during pregnancy but there was no significant association found between baseline BMI and ESS scores. This study did not note any weight changes during pregnancy. (19)

For formal diagnosis, the gold standard is overnight, observed 24-hour polysomnography as per American Academy of Sleep Medicine (AASM) guidelines. (2,20) A cheaper and technically simpler alternative is a wrist-worn monitor such as Watch-Pat 200. (17) The RUSleeping monitor (Philips Respironics, The Netherlands)
used in this study is worn on the face and monitors changes in nasal pressure to detect respiratory events. (21) It has been classified as a single-channel ASDA level IV device and gives an hourly and cumulative AHI score. A study of 25 adults with suspected OSA demonstrated good agreement between data collected by the RUS meter and a standard multi-channel polysomnogram (R=0.77, p<0.001.) with high sensitivity (89%) and specificity (86%). (21) However, there is no previous data on the use of the RUS meter during pregnancy.

This study aims to investigate the prevalence of OSA in obese, pregnant woman using the ESS questionnaire to identify those with symptoms and the RUS meter to provide definitive data. Accurate identification of those with OSA will help ensure correct treatment and management to help reduce rates of complications.

**Materials and methods**

The Fit for Birth study was a cohort study of pregnant women with a BMI of ≥30. Detailed information is available elsewhere (Narayanan et al. 2016) but in brief 824 women were recruited over a one-year period in a large tertiary referral hospital in the northwest of England. (22) Data was collected for all those in the study who subsequently gave birth at the trust.

For this nested study on OSA in pregnancy (the ‘Fit for Birth PLUS’ study), additional data was collected from 162 women with BMI ≥35. These women attended a research antenatal clinic 3 times during their pregnancy, and changes in weight during pregnancy were recorded along with data on diet, activity, sleep and quality of life. Visits were made at approximately 16 weeks, 28 weeks and 36 weeks gestation.
This paper focuses on the sleep data that was collected using Epworth Sleepiness Scale (ESS) questionnaires completed at each trimester visit plus data collected by the RUS meter offered at approximately 28 weeks gestation. Ethical approval was obtained from Liverpool (Adult) Research Ethics Committee (09/H1005/23) and each woman gave individual informed consent for participation.

The Epworth Sleepiness Scale is a standardised questionnaire that measures the probability of falling asleep in eight situations. (23) Candidates rank the likelihood from 0 to 3, with 0 being extremely unlikely to fall asleep and 3 being very likely. The maximum score is 24 and \( \geq 10 \) is considered abnormal, requiring further investigation. It has previously been validated for use in the general population and has been used in other studies on pregnant women. (24)

The RUSleeping RTS meter (Philips Respironics, The Netherlands) is a screening device worn on the face that monitors changes in nasal pressure to detect respiratory events. At approximately 28 weeks, study participants were asked to use the RUSleeping meter overnight at home as an objective and instrumental means of measuring apnoea and hypnoea events in the least intrusive way available at the time. An AHI \( \geq 15 \) was used as an indication for further investigation.

**Data Analysis**

All data taken from consent forms, hospital notes and the hospital’s computer records (MEDITECH®) were collated and anonymised prior to statistical analysis. Analyses were carried out in Intercooled Stata11 (Statcorp, College Station, TX, USA). SPSS was used to analyse change in Epworth Sleepiness Scale scores throughout
pregnancy. The random effects model was used to model change in Epworth scores over time.

**Results**

Data from 162 women was analysed. Their mean age was 29.6 (range 18-43, SD 6.1) and median BMI was 38 (range 35-68.6; Table I). Full details of the cohort are published elsewhere. (22)

Of the 162 who joined the nested ‘Fit for Birth PLUS’ study, 77 women (47.5%) had a complete data set having attended all 3 clinics and having completed the ESS at each of these. Only 47 women completed the use of the RUS meter. (Figure I)

The overall rate of women with an ESS score of ≥ 10 was 12.1% with increasing rates through pregnancy (table II).

Women who were smokers at the time of booking reported slightly higher scores (+1.7) but this is statistically marginal (p=0.04). Several other covariates were investigated and found not to be significant including BMI, weight change, baby’s birthweight and hypertension.

Further descriptive analysis was undertaken to examine individual ESS questions whose findings demonstrate inconsistency within the overall questionnaire results (table III).
A total of 77 women completed ESS questionnaires at all three clinic visits. A repeated measures ANOVA with a Greenhouse-Geisser correction found that the mean Epworth score increased with gestation (M=4.67, M=5.44, M=5.86). This was statistically significant between trimester one and two and between trimester one and three (p=0.009). When comparing solely data from women who visited clinic one and three, the average change in Epworth score was 0.95 (SD=3.7).

There was substantial individual variation between the women in both average score and trend over time. Some increased in score over time, some scores remained flat and some decreased. A random effects model confirmed that the overall scores tended to increase by about 0.058 per week of gestation (approx. 1.2 over 20 weeks gestation) and that the overall average scores and the rates of change varied from woman to woman. In general, the average score was negatively correlated to the direction of change; those with the higher average scores were those with scores that tended to decrease over the visits.

The uptake of women prepared to use the RUS meter was poor. Of the 47 who did use the meter, only one had an AHI index of ≥15. The reasons given for not using the meter are given in table IV.

Out of a possible 125 women who attended this second visit clinic, 47 (37.6%) took the meter home, used it and returned the results. Only 1 of 47 (2.1%) had a score ≥15 which demonstrated a preliminary diagnosis and need for further investigation and likely intervention. This participant had an AHI score of 42.6, which corresponds to the severe OSA category (score ≥30). She attended all three clinics and has ESS
scores of 5, 10 and 21. She had numerous risk factors for and associations with OSA including increased maternal age (39), BMI ≥ 35 and a coexisting mental health disorder. During the time between clinics two and three, she was admitted to the hospital due to raised BP and proteinuria although later discharged. She was eventually induced at 38 weeks gestation and delivered a healthy baby by emergency caesarean section for fetal distress.

**Discussion**

In this study of 162 obese pregnant women with a BMI >35, there was a 12.1% rate of OSA as determined by an ESS questionnaire score of ≥10. This risk increased significantly with each trimester, but did not correlate to booking BMI, weight change, baby birthweight or hypertension. Of those using the RUS meter, the rate of OSA was 2.1%.

The findings of this study are complicated by factors that may cause ESS scores to rise during pregnancy irrespective of OSA. ESS scores may be falsely highly reported particularly in women due to the difference of symptom presentation. ESS scores may also be under-reported in the pregnant population due to commitments with caring for young families and work commitments. The negative correlation between average score and direction of ESS score change might be one aspect that could have been influenced by this confounding factor. There is also an accepted belief that poor sleep and excessive daytime somnolence is part of a normal pregnancy. (4)

NICE recommend the ESS questionnaire as a screening tool for the non-pregnant population, with referral advised if symptomatic plus ESS score >10. (25) However
there are no studies that specifically assess the ESS questionnaires on the pregnant, obese population. Comparative data is therefore lacking.

Responses to individual questions within the ESS were of interest. As expected, nearly no-one reported a “Likelihood of dozing when sitting & talking to someone” or “Likelihood of dozing when in a car, stopped at traffic lights”. For others, answers were very variable. For example, “Likelihood of dozing when lying down to rest in afternoon” had an even spread of results where it might have been assumed that many pregnant women given the chance, would have easily dozed at this point. Many women voiced the view that they would never get chance to lie down in the afternoon due to their other offspring needing attention. These pregnancy specific responses suggest that the ESS may not be as robust during pregnancy. Baumgartel et al take the view that the ESS should be scored using two parts: sleepiness in appropriate and inappropriate situations. High scores of sleepiness in inappropriate situations could indicate higher risks of further complications in pregnancy. (26)

In this study, a score of ≥15 using the RUS meter was used to preliminarily diagnose OSA. This device is a relatively untested method of assessing OSA and had a high rate of non-compliance. Although thorough education was provided on ideal techniques for use, the sleep was unobserved so it is difficult to ascertain the validity of the results achieved. In 2009, RUS meter was considered the most appropriate and cost-effective device to measure apnoeic/hypopnoeic episodes, however, with advancing technology, future studies could achieve better compliance with a newer, less intrusive wrist worn device, for example the Watch-Pat 200. (17) Gold standard investigation as defined by AASM guidelines to diagnose OSA is overnight observed
polysomnography testing, however this is an impractical and expensive alternative. (1)

Of the 47 participants using the RUS meter, one had an AHI score of ≥15 (2%). Interestingly, her score was 42.6 placing her in the severe OSA category. She had a complicated pregnancy and delivery including investigations for pre-eclampsia and the need for an emergency C-section due to fetal distress. These complications have all been found to be associated with OSA. (13-15) There are no other studies using the RUS meter in pregnancy. However, a study of 25 non-pregnant adults referred to a sleep centre with OSA found 18 participants (72%) had an AHI ≥5 and only 5 (25%) had an AHI ≥30. (21) It is difficult to compare these two studies directly and ultimately due to substantial technological advances over recent years the use of the RUS meter has been largely replaced by superior alternatives.

There were a number of weaknesses identified following the implementation of this study. There was a high rate of attrition, with reasons not documented and unable to be ascertained. There is no evidence that women with high Epworth scores were less likely to attend the research clinics, although it is plausible that women with higher scores will be more tired during the day and hence less able to make appointments. There was also a high rate of non-compliance with the RUS meter, making it difficult to link results from the ESS questionnaire to the RUS meter results.

In conclusion, this study has found that 12% of obese pregnant women scored positive on the ESS, but this was not confirmed on formal testing using more invasive techniques. Although OSA in pregnancy is important, this study identified problems
with both the ESS and the RUS meter and neither appeared to provide a reliable method with which to diagnose OSA in pregnancy. Larger studies are needed before a reliable estimate of the prevalence of OSA in pregnancy is made.

**Conflicts of interest statement**

The authors declare that they have no conflicts of interest.

**Acknowledgements**

HL collected data and wrote the paper with KM. AW and JW provided input and edited article. All authors were involved with the main Fit For Birth study and approved final submitted paper.

The authors would like to thank the Liverpool Primary Care Trust for funding this project, Michelle Dower (research midwife) and Dr Joyce Carter for her valuable inputs during the study. We would also like to acknowledge the contribution of community midwives who participated in this study.

**Funding**

Funded by Liverpool Primary Care Trust. Donation of RUSleeping devices from Philips/Respironics.

**References**


### Tables

Table I: BMI of the study participants

<table>
<thead>
<tr>
<th>BMI category</th>
<th>Total number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>35-39.9</td>
<td>94 (58%)</td>
</tr>
<tr>
<td>40-44.9</td>
<td>41 (25.3%)</td>
</tr>
<tr>
<td>45+</td>
<td>27 (16.7%)</td>
</tr>
</tbody>
</table>

Table II: Summary ESS scores

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Mean score (SD; range)</th>
<th>ESS ≥ 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; visit – 16 weeks</td>
<td>143</td>
<td>4.82 (3.66; 0-19)</td>
<td>10 (7.0%)</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; visit – 28 weeks</td>
<td>125</td>
<td>5.30 (3.85; 0-20)</td>
<td>15 (12.0%)</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; visit – 36 weeks</td>
<td>93</td>
<td>5.87 (4.10; 0-21)</td>
<td>16 (17.2%)</td>
</tr>
</tbody>
</table>
Table III: Responses to selected questions within the ESS.

“Likelihood of dozing when sitting & talking to someone”

<table>
<thead>
<tr>
<th>Visit</th>
<th>Score 0 (% of total responders)</th>
<th>Score 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visit 1 (n = 143)</td>
<td>N=135 (94.4%)</td>
<td>N=1 (0.7%)</td>
</tr>
<tr>
<td>Visit 2 (n = 125)</td>
<td>N=117 (94.3%)</td>
<td>N=0</td>
</tr>
<tr>
<td>Visit 3 (n = 93)</td>
<td>N=83 (91.2%)</td>
<td>N=0</td>
</tr>
</tbody>
</table>

“Likelihood of dozing when in a car, stopped at traffic lights”

<table>
<thead>
<tr>
<th>Visit</th>
<th>Score 0 (% of total responders)</th>
<th>Score 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visit 1 (n = 143)</td>
<td>N=137 (95.8%)</td>
<td>N=0</td>
</tr>
<tr>
<td>Visit 2 (n = 125)</td>
<td>N=117 (94.3%)</td>
<td>N=1 (0.8%)</td>
</tr>
<tr>
<td>Visit 3 (n = 93)</td>
<td>N=84 (92.3%)</td>
<td>N=0</td>
</tr>
</tbody>
</table>

“Likelihood of dozing when lying down to rest in afternoon”

<table>
<thead>
<tr>
<th>Visit</th>
<th>Score 0 (% of total responders)</th>
<th>Score 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visit 1 (n = 143)</td>
<td>N=11 (7.4%)</td>
<td>N=46 (31.1%)</td>
</tr>
<tr>
<td>Visit 2 (n = 125)</td>
<td>N=16 (12.9%)</td>
<td>N=38 (30.6%)</td>
</tr>
<tr>
<td>Visit 3 (n = 93)</td>
<td>N=9 (9.9%)</td>
<td>N=28 (28%)</td>
</tr>
</tbody>
</table>
Table IV: Reason for non-compliance with RUS meter

<table>
<thead>
<tr>
<th>Reason for non-compliance</th>
<th>Total number</th>
<th>Percentage of total non-compliant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declined (various reasons: anxiety, unwell, didn’t see the point, didn’t like the idea, not interested etc.)</td>
<td>41</td>
<td>25.3%</td>
</tr>
<tr>
<td>No meters available to give out</td>
<td>13</td>
<td>8.0%</td>
</tr>
<tr>
<td>Didn’t return/returned not used</td>
<td>12</td>
<td>7.4%</td>
</tr>
<tr>
<td>Not asked to take meter by satellite clinic</td>
<td>9</td>
<td>5.6%</td>
</tr>
<tr>
<td>Not asked due to language barrier</td>
<td>1</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

Figures

Figure I: Numbers of participating women attending each clinic

- Women in fit for birth PLUS subset (n=162)
  - Attending clinic at 16 weeks and completing ESS (n=143) 86.6%
  - Attending clinic at 28 weeks and completing ESS (n=125) 77.2%
  - Attending clinic at 36 weeks and completing ESS (n=93) 57.4%
- Uptake of RUS meter (n=47) 29.0%