Introduction

For the purpose of this report, health data was examined from a simple random sample of adults in the United States. Each adult was asked a series of health-related questions, and the data compiled into a 2012 BRFSS survey. The survey questions most relevant to this report included sex, marital status (married vs. never married), BMI (body mass index; an indicator of body fatness), and amount of sleep obtained per day (average number of hours sleep obtained in a 24-hour period). The following 3 questions were asked:

1) Do married women tend to be heavier (as measured by BMI) than never married women?
2) Do married men tend to be heavier (as measured by BMI) than never married men?
3) Do married men tend to sleep more than married women?

Question 1 - Do married women tend to be heavier (as measured by BMI) than never married women?

In order to address this question, I asked "Is there a difference between the mean BMI of married women and the mean BMI of never married women?"

Figures 1 and 2 display histograms with data that is not normally distributed. Both histograms show moderate skewness to the right. The standard deviation for married women (5.7 units) is similar to the standard deviation for never married women (6.4 units). The mean BMI of the 528 married women (27.2 units) is also similar to the mean BMI of the 129 never married women (26.9 units).

Figure 1 Histogram of BMIs for 528 married women

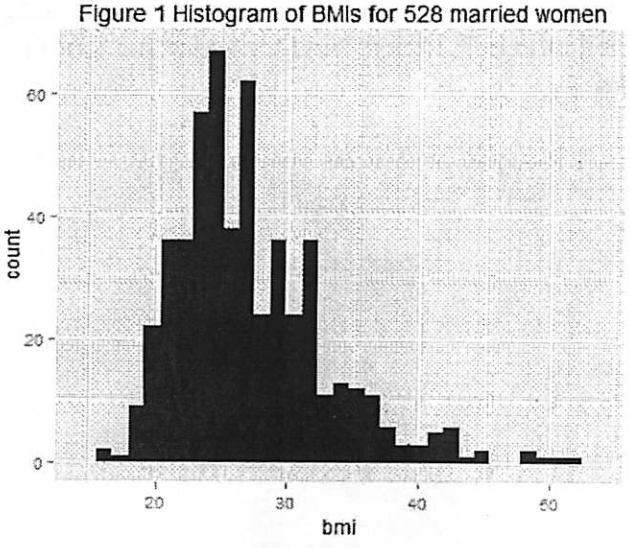
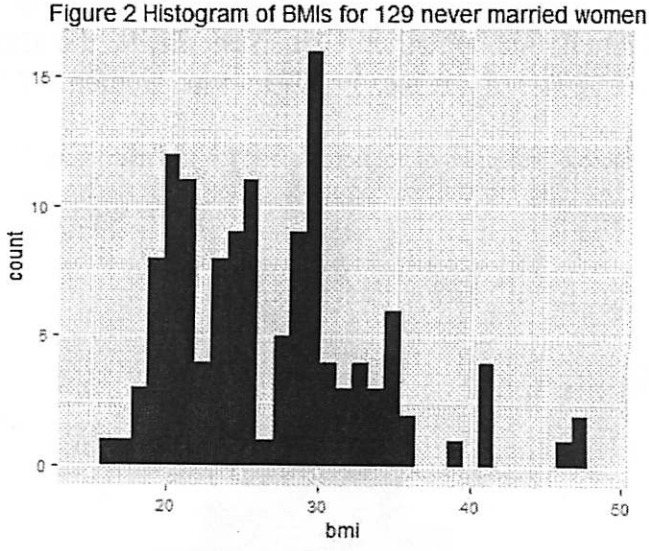


Figure 2 Histogram of BMIs for 129 never married women



Question 2 - Do married men tend to be heavier (as measured by BMI) than never married men?

To address this question, I asked "Is there a difference between the mean BMI of married men and the mean BMI of never married men?"

Figures 3 and 4 display histograms with data that is relatively normally distributed. The two histograms have a relatively similar shape, with mild skewness to the right. The standard deviation for married men (4.9 units) is very similar to the standard deviation for never married men (4.8 units). The mean BMI of the 426 married men is 28.1 units, and the mean BMI of the 147 never married men is 26.0 units.

Figure 3 Histogram of BMIs for 426 married men

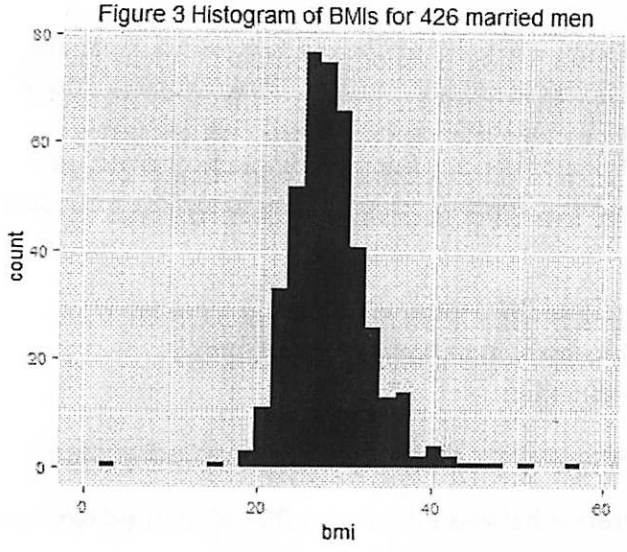
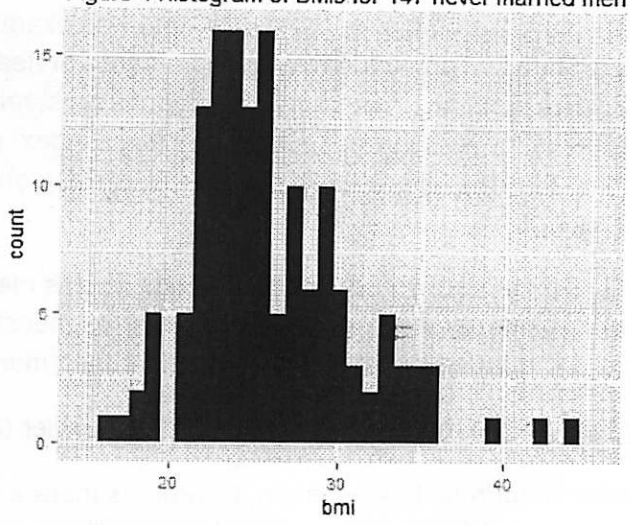


Figure 4 Histogram of BMIs for 147 never married men



Question 3 - Do married men tend to sleep more than married women?

In order to address this question, I asked "Is there a difference between the mean number of hours sleep per day for married men and the mean number of hours sleep per day for married women?"

Figures 5 and 6 display histograms with data that is relatively normally distributed. The two histograms have a similar shape. The standard deviation is the same for married men and married women (1.3 hours). The mean number of hours sleep per day for the 428 married men is 7.2 hours, and the mean number of hours sleep per day for the 562 married women is 7.0 hours.

Figure 5 Histogram of no. hours sleep for 428 married men

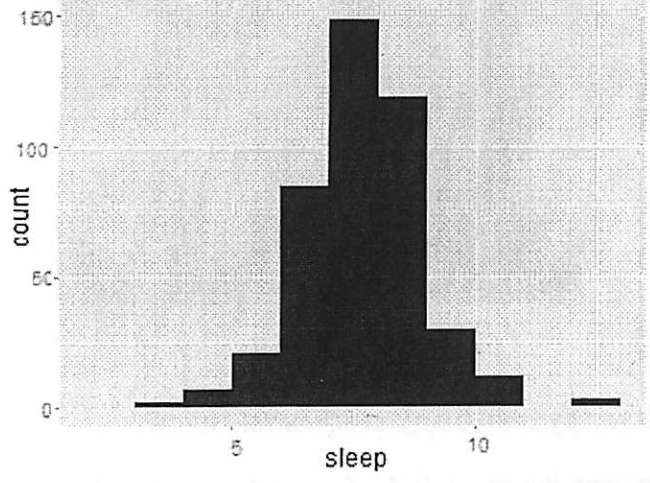
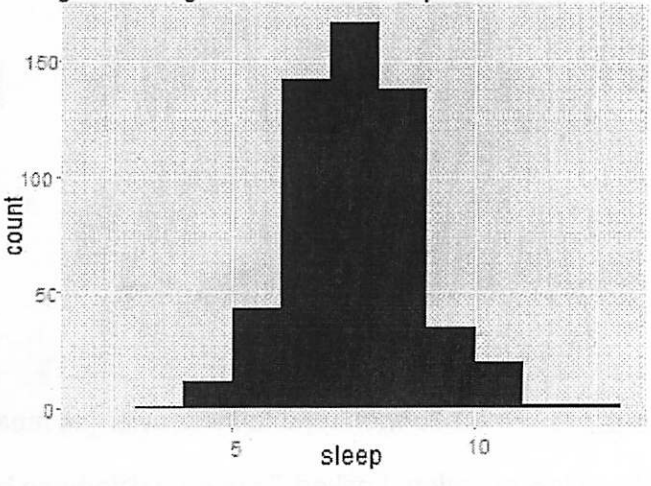


Figure 6 Histogram of no. hours sleep for 562 married women



Methods

A two sample t-test was utilized to answer questions 1, 2 and 3. Assumptions of the two sample t-test include normal populations, equal population standard deviations, and independence of observations within and between groups. I am confident that the final assumption has been met for all 3 cases, as it has been outlined that all data comes from a simple random sample of adults in the USA.

Question 1 - Is there a difference between the mean BMI of married women and the mean BMI of never married women?

A two sample t-test was utilized because our question of interest is about population means, specifically comparing the difference between the mean of one population (BMI of married women) and the mean of a second population (BMI of never married women) that has been independently sampled. The standard

deviations of the two groups are similar (5.7 units for married women, 6.4 units for never married women) suggesting that the equal population standard deviation assumption has been met. Histograms of the data, however, are not normally distributed and both samples show moderate skewness to the right (see Figures 1 and 2). The normal probability plots for the samples also fall far from the straight line (see Figure 7). Skewness would substantially affect validity of the t-tools for this case, since the sample sizes differ considerably (528 married women, 129 never married women). Although the sample sizes are quite large giving robustness to the violation, I felt that validity of the test would be improved by log transforming the data. The goal of a log transformation was to reduce the level of skewness seen in both samples and thereby improve validity of the t-test.

Figure 7 Normal probability plots for BMIs for women

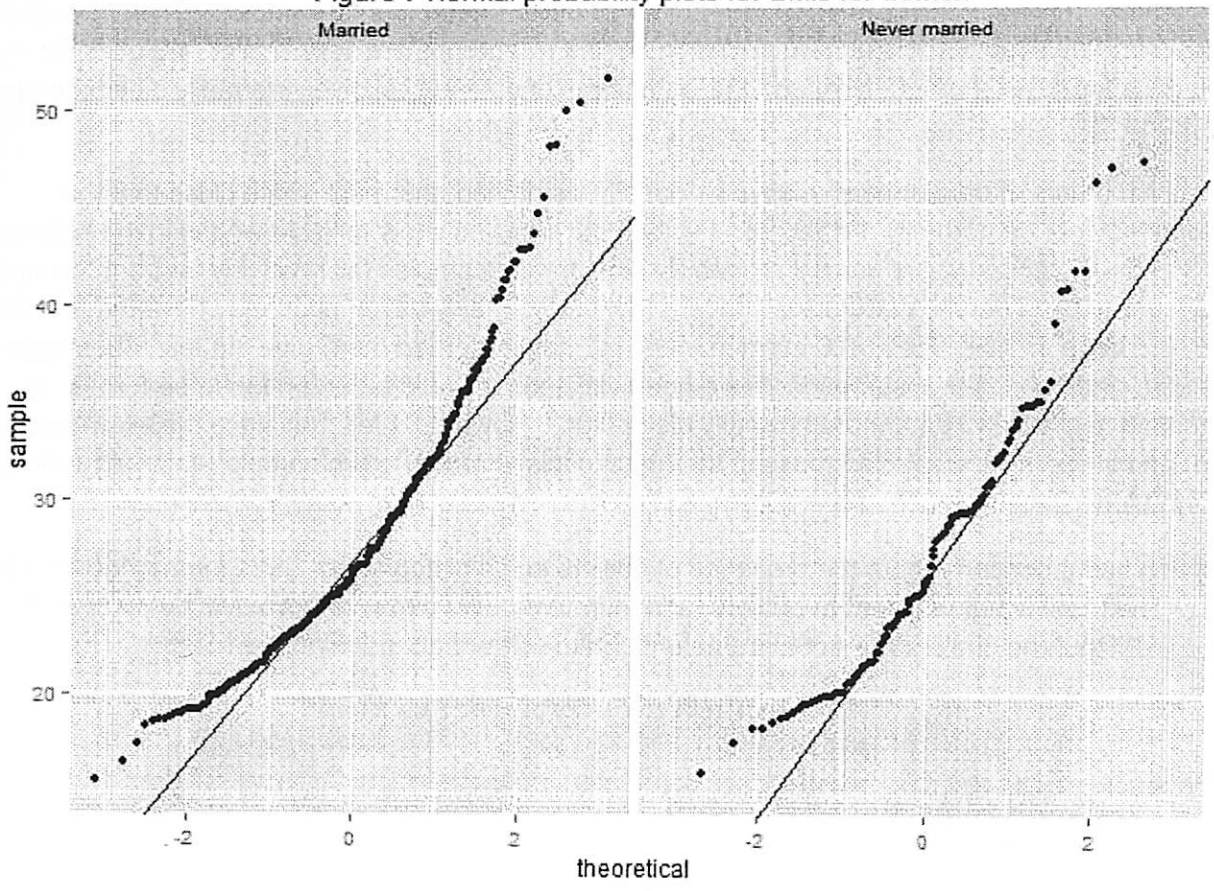


Figure 8 Histogram of log BMIs for 528 married women

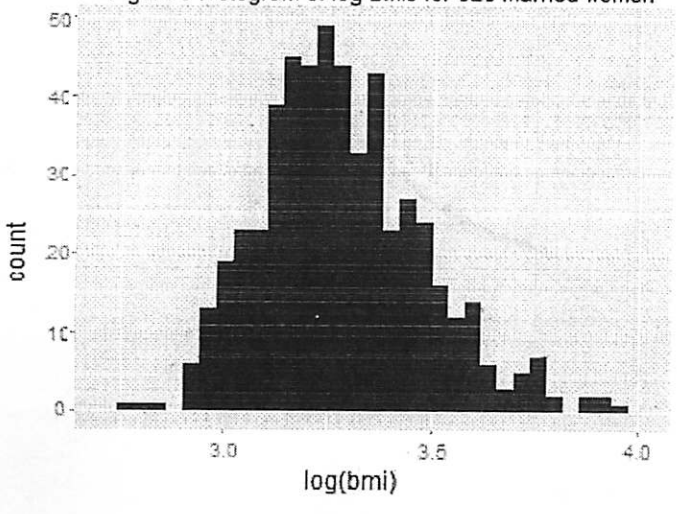
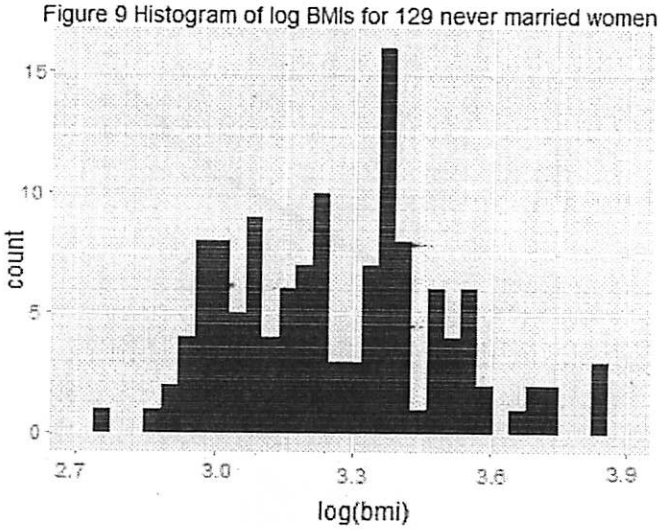


Figure 9 Histogram of log BMIs for 129 never married women



Histograms of the log transformed data are now normally distributed, with substantially less skewness. The log transformed BMIs meet the assumption of normally distributed data. The standard deviation of the log BMI for married women (0.20 units) is very similar to the standard deviation of the log BMI for never married women (0.23 units). A two sample t-test on the log transformed data is therefore appropriate as the assumptions of normal populations, equal standard deviations, and independence have been met.

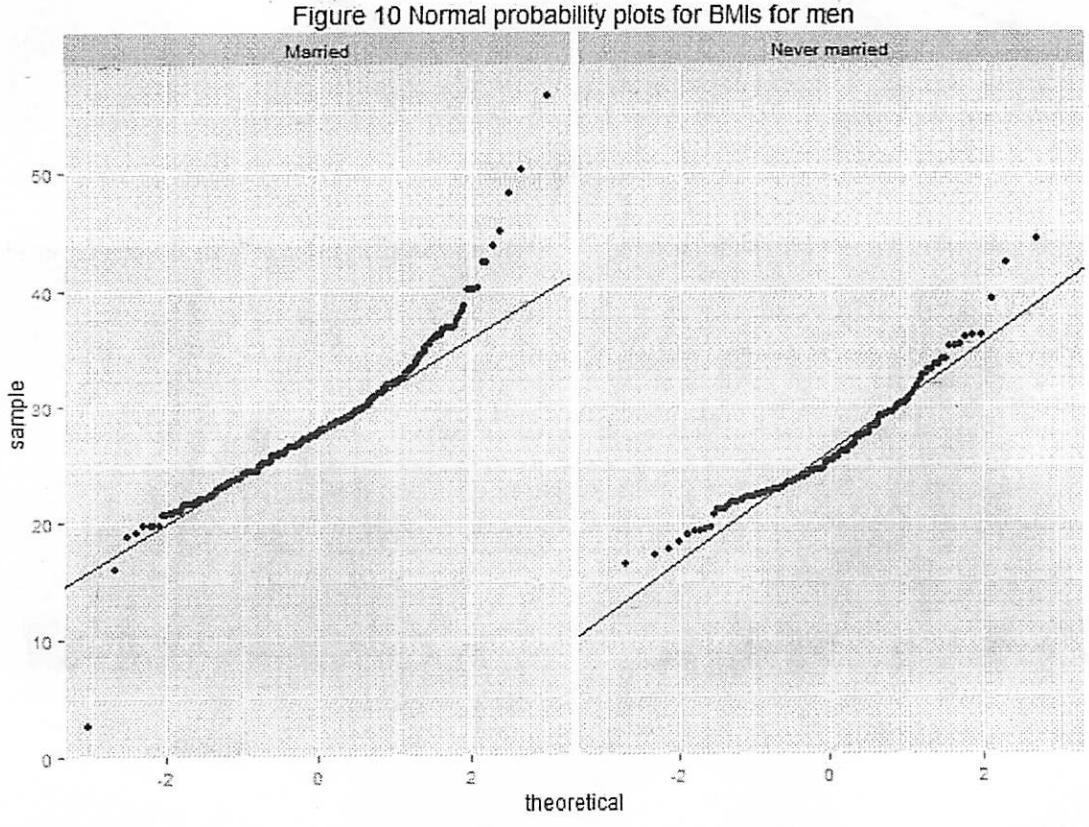
Question 2 - Is there a difference between the mean BMI of married men and the mean BMI of never married men?

A two sample t-test was utilized because our question of interest is about population means, specifically comparing the difference between the mean of one population (BMI of married men) and the mean of a second population (BMI of never married men) that has been independently sampled. The two sample t-test is appropriate because the standard deviations of the two groups are very similar (4.9 units for married men, 4.8 units for never married men) and the histograms of data are relatively normally distributed (see Figures 3 and 4).

While the histograms of data appear relatively normally distributed, there are some clues that the populations may not be normally distributed. Firstly, both samples show mild skewness to the right (see Figures 3 and 4). Secondly, the normal probability plots for the samples fall relatively far from the straight line (see Figure 10). There is therefore a concern that the normality assumption may have been violated. However, the large sample sizes (426 married men, 147 never married men) give the test robustness to this violation, since no other assumptions have been violated. It would be an option to perform a log transformation to make the data more normally distributed. However, I felt that since the skewness was only mild, and since there was robustness of the t-test to this violation, a two sample t-test was valid without a transformation.

There is also some concern about the presence of outliers in the histogram of BMIs for married men (see Figure 3). The t-test is not resistant to outliers, and therefore outliers may affect validity of the results. However, I feel that the outliers are not extreme enough to warrant abandoning the t-test.

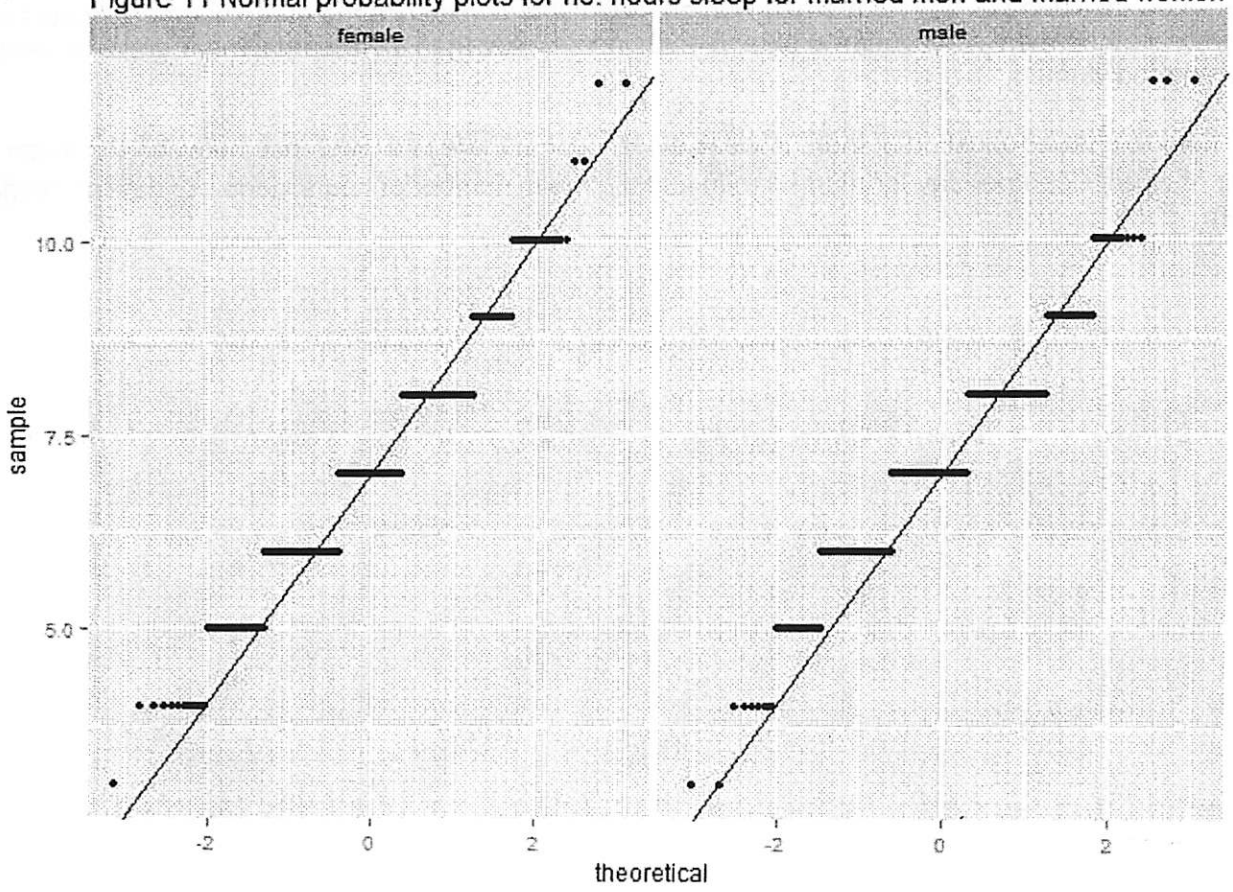
Figure 10 Normal probability plots for BMIs for men



Question 3 - Is there a difference between the mean number of hours sleep per day for married men and the mean number of hours sleep per day for married women?

A two sample t-test was utilized because our question of interest is about population means, specifically comparing the difference between the mean of one population (number of hours sleep per day for married men) and the mean of a second population (number of hours sleep per day for married women) that has been independently sampled. The two sample t-test is appropriate because the standard deviations of the two groups are equal (1.3 hours) and the histograms of data appear normally distributed (see Figures 5 and 6). The normal probability plots for the samples also fall close to the straight line (see Figure 11).

Figure 11 Normal probability plots for no. hours sleep for married men and married women



Summary

Question 1 - Is there a difference between the mean BMI of married women and the mean BMI of never married women?

There is no evidence that the mean log BMI of married women is not equal to the mean log BMI of never married women (two sample t-test, two-sided p-value = 0.36). We estimate that the median BMI of married women is 1.02 times the median BMI of never married women. With 95% confidence the median BMI of married women is between 1.0 and 1.06 times the median BMI of never married women.

In summary to answer question 1, there is no evidence that there is a difference between the mean log BMI of married women and the mean log BMI of never married women.

Question 2 - Is there a difference between the mean BMI of married men and the mean BMI of never married men?

There is convincing evidence that the mean BMI of married men is not equal to the mean BMI of never married men (two sample t-test, two-sided p-value < 0.001). The mean BMI of married men is estimated to

be 2.1 units higher than the mean BMI of never married men. With 95% confidence the mean BMI of married men is between 1.2 and 3.0 units higher than the mean BMI of never married men.

In summary to answer question 2, there is convincing evidence that the mean BMI of married men is different to the mean BMI of never married men.

Question 3 - Is there a difference between the mean number of hours sleep per day for married men and the mean number of hours sleep per day for married women?

There is weak evidence that the mean number of hours sleep per day for married men is not equal to the mean number of hours sleep per day for married women (two sample t-test, two-sided p-value = 0.09).The mean number of hours sleep per day for married men is estimated to be 0.14 hours higher than the mean number of hours sleep per day for married women. With 95% confidence the mean number of hours sleep per day for married men is between - 0.02 and 0.30 hours higher than the mean number of hours sleep per day for married women.

In summary to answer question 3, there is weak evidence that there is a difference between the mean number of hours sleep per day for married men and the mean number of hours sleep per day for married women.