MAS6061 Project 1 Epidemiology 2017/18

Students should send answers in Word or PDF documents. This assignment should be submitted via MOLE on Tuesday 14th November by 12:00. **These exercises form part of the assessment of the module (Total marks of 60, worth 30% of the mark for Epidemiology).**

1. The number of heart-beating cadaver donors for 2000, broken down into 10 year age bands was obtained from NHS Blood and Transplant (NHSBT). The table below shows the age distribution of these donors for the United Kingdom as a whole and for Scotland, with the corresponding population age distributions at the 1991 decennial census.

Age Band	UK Donors	UK Population	Scotland Donors	Scotland Population
0-9	12	7,395,033	1	610,385
10-19	78	7,524,570	8	647,952
20-29	83	7,677,822	11	662,645
30-39	90	9,524,953	6	818,884
40-49	165	7,895,663	12	707,375
50-59	206	7,262,395	17	620,934
60-69	90	5,429,785	6	493,845
70-79	14	4,293,963	1	369,438

(i) For the United Kingdom calculate the age specific organ donation rates per 1,000,000 for the whole period. **[3 marks]**

(ii) Calculate the Standardised Donation (Incidence) Rate (SDR) for organ donations for Scotland.

[10 marks]

(iii) Calculate the 95% confidence interval for this SDR using the log transformed standard error**.**

[5 marks]

(iv) Does the number of organ donations in Scotland appear particularly high compared to the UK?

[2 marks]

2. A large population study is conducted to assess the association between obesity and cardiovascular disease in secondary school teachers. Teachers aged between 35 and 55 were recruited to the study and followed up for ten years. In this study obesity was defined as BMI greater than or equal to 30 at baseline recruitment. Cardiovascular disease (CVD) risk is known to increase with age, so age may be an important confounding variable. The results of the study are shown below.

	Age <= 45		**Age > 45**	
	CVD	No CVD	CVD	No CVD
BMI >= 30	105	450	240	600
BMI < 30	120	1800	60	900
	225	2250	300	1500

(i) What type of study design is this? According to the ***classic definition of confounding***, is there evidence of confounding in the relationship between BMI and being diagnosed with CVD?

[4 marks]

(ii) Calculate the difference in risks *(Risk Difference)* for cardiovascular disease by BMI exposure for the subgroup aged 45 and below at study onset, the subgroup aged over 45 and the collapsed (over age group) table i.e. three risk differences in total.

[6 marks]

(iii) Comment on whether there is evidence of confounding according to the collapsibility definition. **[2 marks]**

(iv) Compare your answers from sections i) to iii). What is the overall evidence that age is a confounder and/or an effect modifier in the relationship between BMI and CVD in secondary school teachers? How are the results of this association best reported (i.e. should crude, adjusted or strata specific risk differences be reported)?

[3 marks]

3. The table below gives the results of four case control studies of women in North America examining the relationship of passive smoking exposure to lung cancer.

a) Calculate the Odds Ratios of lung cancer associated with passive smoking exposure for each of the four studies separately. Which study shows the highest odds of having lung cancer associated with passive smoking exposure? **[5 marks]**

b) Use the Mantel-Haenzel procedure to find an estimate and confidence interval for the odds ratio of lung cancer associated with passive smoking exposure. [**10 marks]**

Exposure to passive smoking among female lung cancer cases and controls in four studies

	Lung cancer cases		***Controls***	
Study	***Exposed***	***Unexposed***	***Exposed***	***Unexposed***
1	105	30	277	125
2	32	11	160	36
3	11	13	16	9
4	17	5	78	54

c) Comment on your results. What factors need to be borne in mind when generalising the results from the Mantel Haenzel analysis?

[5 marks]

4. A case control study was carried out to investigate the relationship between physical activity in early pregnancy and the risk of pre-term birth in new mothers. Cases were new mothers who had given birth pre-term and controls were new mothers who had given birth full-term. The study comprised three hundred and three matched pairs of women who had given birth at the same hospital during a one year period. The women were matched by age (in years) and parity (number of previous full-term births). The exposure of interest was the mother's predominant work activity in the first four months of pregnancy, standing vs sitting or walking. The outcome of interest was pre-term or full-term birth.

Each woman with a pre-term birth was matched with a woman with a full-term birth and the information on their exposure to 'standing' during the first four months was then measured. Of the 303 pairs of women, 193 were concordant for standing in both pre and full-term births, 63 were standing for pre-term and not standing in full-term, 39 were not standing for pre-term and standing for full-term and 8 were concordant for not standing in both pre and full-term births.

a) Analyse the paired data appropriately (taking account of the matching) and report an odds ratio along with a p-value and 95% confidence interval for the odds of a pre-term birth when exposed to standing compared to a full-term birth when exposed to standing. Comment on the results; is physical activity associated with the risk of a pre-term birth? **[3 marks]**
b) Contrast the results in a) with what you would get (odds ratio, p-value and 95%CI) if you ignored the matching and treated the paired data as unpaired (i.e. with 303 cases and 303 controls). **[2 marks**]
c) Now consider if the original study matched pair data had been slightly different with the same discordant numbers but those concordant for standing were reduced to 143 and those concordant for not standing were 58. Repeat the calculations in a) and b) and comment on the differences when the concordant counts have changed**.** **[5 marks]**