Epidemiology for Health and Medical Sciences II

Sample exam questions - SOLUTIONS

<u>Question A</u>

The following description is adapted from a published abstract.

Aoki et al. Incidence of injury among adolescent soccer players: a comparative study of artificial and natural grass turfs. Clin J Sport Med 2010; 20: 1-7.

The objective of this research was to investigate the incidence of acute injuries and soccer-related chronic pain from long-term training and during matches in adolescent players using natural grass turfs and artificial turfs. Study participants were youth soccer players (12-17 years of age) from 6 teams, with a predominant tendency to train on either natural turf or artificial turf. Of 332 players enrolled in this study, 301 remained to completion. Medically diagnosed acute injuries and chronic pain were recorded daily by team health care staff throughout 2005 and this information was provided to the researchers.

Acute injuries per 1000 player hours on each surface and chronic complaints per 1000 player hours were evaluated according to frequency of surface used most of the time. There was no significant difference in the incidence of acute injuries between the two surfaces during training and competition. However, the artificial turf group showed a significantly higher incidence of low back pain during training (RR = 1.63, 95% confidence interval = 1.06-2.48). Age (early rather than late adolescence) and prolonged training hours were factors associated with an increased incidence of chronic pain in the artificial turf group.

CONCLUSION: Adolescent players routinely training on AT for prolonged periods should be carefully monitored, even on AT conforming to new standards.

a. What specific study design has been used by Aoki and colleagues? **[1 mark]**

A cohort study (prospective).

b. In your own words, describe the relationship between turf type used for training and low back pain. Include an explanation and interpretation of the information contained in the brackets. **[3 marks]**

Those who trained on artificial turf had a risk of low back pain during training that was 62% higher than the risk for those who trained on natural grass. The observed data are consistent with RRs ranging from a 1.06 (6% increase) to 2.48. This interval does not contain RR=1, so data not compatible with groups having same risk – this finding is statistically significant at the traditional 0.05 level.

c. Complete the 2x2 table and show how the relative risk of 1.63 was calculated. **[3 marks]**

Turf	Had low back pain	Did not have low back pain	
Artificial	45	90	135
Natural	34	132	166
	79	222	301

$$RR = \frac{45 / 135}{34 / 166}$$

$$= 45 / 135 \times 166 / 34 = 1.63$$

d. A total of 332 players were enrolled and 301 completed the study. Suppose that almost all of the 31 players who did not complete the study came from teams that trained on artificial turf. Thinking as an epidemiologist, explain why you would be concerned about this and how it might affect the results of the study. **[2 marks]**

An epidemiologist would recognize this as a systematic difference between the groups, possibly creating bias. If the players who did not complete the study were injured, this would mean injuries were under-estimated those who trained on artificial turf. On the other hand, if those players tended to be uninjured, the available data would over-estimate injuries among those who trained on artificial turf. Both of these scenarios would change the RR.

Question B

Read the following adapted abstract of a published paper.

Hurley et al. Tobacco smoking and alcohol consumption as risk factors for glioma: a case-control study in Melbourne, Australia. J Epidemiol Community Health 1996; 50: 442-6.

OBJECTIVE: To investigate possible associations between tobacco smoking and alcohol consumption and the risk of adult glioma.
DESIGN: This was a population based, case-control study. Relative risks (RR) were estimated using logistic regression analysis.
SETTING: Melbourne, Australia.
PARTICIPANTS: These comprised 416 case subjects (166 women, 250 men), 66% of those eligible; and 422 control subjects (170 women, 252 men), 43.5% of those potentially eligible.
RESULTS: There was no increase in risk of glioma with having ever smoked tobacco (RR 1.29, 95% CI 0.95, 1.75) for all subjects, adjusted for age, a reference date, and gender. There was a slight increase in risk for men (RR 1.64, 95% CI 1.1, 2.45), but not for women (RR 0.99, 95% CI 0.62, 1.62). For men, there was no increase in risk with increasing pack-years of cigarette smoking, but the risk was significantly increased in subjects who had smoked for less than 10 years. There was no increase in risk associated with having ever drunk alcohol for all subjects (RR 0.96, 95% CI 0.67, 1.37), women (RR 0.69, 95% CI 0.4, 1.15) or men (RR 1.40, 95% CI 0.81, 2.43).
CONCLUSIONS: This study does not support an association between either tobacco smoking or alcohol consumption and glioma. The pattern of risk associated with tobacco smoking in men appears inconsistent with a causal role, and may be due to chance, response bias, or uncontrolled confounding.

[Note: remember that in a case-control study an odds ratio is what is actually calculated. But this provides an estimate of RR where a disease is rare. The above is an example of authors using that principle.]

a. Glioma is the most common type of primary brain tumor in adults. Glioma tends to kill people or leave them mentally incapacitated. Where cases had died or were unable to complete questionnaires, information was obtained from a relative or friend.
Explain how this might contribute to information bias. What impact would this be likely to have on the relative risks? **[3 marks]**

Relative or friends of cases are not likely to know about smoking and drinking habits in as much detail as cases themselves. They may also want to present the case individuals in the most favourable light. This could lead to systematic underestimates of the amount that the cases smoked or drank.
This would tend to bias estimates of relative risks – would be under-estimates.

b. Controls were persons living in the same suburb as the case, identified through the electoral roll. Only 44% of potential controls participated in the study. Explain how this could create bias. **[2 marks]**

Example. Perhaps the 44% of controls who were willing to participate tended to be health conscious. They might then smoke and drink less than the 66% of people approached who did not want to participate in the control group. This could artificially lead an association between smoking and drinking and the occurrence of glioma.

c. In this study, 10 of the cases and 6 of the controls had a history of meningitis. Create a 2x2 table and calculate the odds ratio for the association between history of mengitis and glioma. **[3 marks]**

	Glioma case	Control
History of meningitis	10	6
No history of meningitis	406	416
	416	422

OR = 10 : 406 / 6 : 416 = 10 / 406 x 416 / 6 = 1.7

<u>Question C</u>

A study from the University of Texas examined whether the risk of Hepatitis C (Hep C) was related to whether people had tattoos. A sample of 600 individuals was randomly selected from the student population. In this sample 113 had a tattoo. For individuals with a tattoo, 22 were found to have Hepatitis C and for individuals without a tattoo, 25 were found to have Hepatitis C. Researchers want to know if there is evidence from this study of an increased prevalence of Hep C for individuals with a tattoo amongst the population from which they were sampled.

a. What study design is being used in this example? **[1 mark]**

A cross-sectional study or survey.

b. Give formal statements of the Null and Alternative hypotheses. **[2 marks]**

H_0: there is no association in the population between having Hep C and having a tattoo.

H_a: there is an association in the population between having Hep C and having a tattoo.

Note: This could also be stated in terms of proportions, i.e.

H_0: in the population the proportion (prevalence) of individuals with Hep C is the same for those with a tattoo as it is for those without.

H_a: in the population the proportion (prevalence) of individuals with Hep C is NOT the same for those with a tattoo as it is for those without.

c. Construct an appropriately labelled 2x2 table to display these data. **[2 marks]**

	Hep C	No Hep C	
Tattoo	22	91	113
No tattoo	25	462	487
	47	553	600

d. Calculate the sample prevalence (risk) ratio for Hep C among individuals with a tattoo compared to individuals without a tattoo. Interpret it. **[3 marks]**

prevalence of HEPC for those with a tattoo = 22/113 = 0.1947

prevalence of HEPC for those without a tattoo = 25/487 = 0.0513

prevalence risk ratio = 0.1947/0.0513 = 3.795 = 3.8 (to one decimal place)

Individuals with a tattoo in this study are 3.8 times more likely to have Hep C than individuals without a tattoo.

e. The formula for a χ^2 test statistic is

$$\chi^2_{stat} = \sum_{all\ cells} \left[\frac{(O-E)^2}{E}\right]$$

Using this formula, the χ^2 test statistic for this study is 26.1 with a p-value of less than 0.001. Show how to use this knowledge to make a decision with respect to your opposing hypotheses and explain the reason for that decision.
[Assume the standard level for statistical significance is 0.05.] **[3 marks]**

Given that Prob ($\chi^2 > 26.1$) <0.001, we can confidently reject the null hypothesis stated in (b) and state that we have very strong evidence (according to Bland) for the alternative hypothesis which states that there is an association in the population between having Hep C and having a tattoo and that the observed association is unlikely to have arisen by chance alone.

Question D

Studies have linked brain size in infants and toddlers to a number of future prognoses including autism. One study looked at the brain sizes of 30 boys diagnosed with autism and 12 boys without autism (controls) who all received an MRI scan as toddlers.

A summary of the results on whole-brain volumes in millilitres[3] can be seen below.

Group	n	mean	standard dev
Autistic	30	1297.6	88.4
Control	12	1179.3	70.7

a. If interest was solely whether there was a difference in whole-brain volumes for infants and toddlers, state an appropriate null and alternative hypothesis for this research question.

[2 marks]

H_0: there is no difference in the infant/toddler *population* between mean whole-brain volume in autistic versus non-autistic toddlers.

H_a: there *is* a difference in the infant/toddler *population* between mean whole-brain volume in autistic versus non-autistic toddlers.

b. An appropriate statistical test was performed in Stata with output seen below. Using Bland's criteria for significance, explain what decision you would make with respect to your hypotheses and what this means in the context of this study. **[3 marks]**

```
ttesti 30 1297.6 88.4 12 1179.3 70.7

Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      30      1297.6    16.13956        88.4    1264.591    1330.609
       y |      12      1179.3    20.40933        70.7    1134.379    1224.221
---------+--------------------------------------------------------------------
combined |      42      1263.8    15.27072    98.96561     1232.96     1294.64
---------+--------------------------------------------------------------------
    diff |               118.3     28.6592                 60.3776    176.2224
------------------------------------------------------------------------------
    diff = mean(x) - mean(y)                                      t =   4.1278
Ho: diff = 0                                     degrees of freedom =       40

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9999         Pr(|T| > |t|) = 0.0002          Pr(T > t) = 0.0001
```

The 95% confidence interval is (60.3776 , 176.2224). In words this means we are 95% confident that the population difference in means lies somewhere between 60.3776 and 176.2224 millilitres[3]. As zero is not contained in this interval we have sufficient evidence against the null hypothesis. As a result, we have evidence that there is a difference in the infant/toddler population between mean whole-brain volume in autistic versus non-autistic toddlers. The evidence (according to Bland) also happens to be 'very strong' as it is less than 0.001.