Assignment advanced epidemiology

## Preamble

This assignment is designed to test your understanding of contemporary epidemiology. It does so by asking you to interpret a seminal paper which is now around 70 years old. We chose an older paper because it discusses epidemiological concepts without using technical language. It therefore allows you to demonstrate an understanding of epidemiological concepts, even where epidemiological language is not being used.

Read all of the article “Smoking and carcinoma of the Lung” which is available at (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2038856/pdf/brmedj03566-0003.pdf>) then answer these questions in short note form.

1. In the introductory section (ie before the heading “Possible Causes of the Increase”), the authors report that:-

“*in the quarter of a century between 1922 and 1947 the annual number of [lung cancer] deaths recorded increased from 612 to 9,287, or roughly fifteenfold*”.

The authors then consider three explanations for this observation; i) increases in the population, ii) increases in the numbers of older people in the population, and iii) improved standards of diagnosis. What epidemiological problem/concepts do each of these explanations illustrate? For each, give a further illustrating example (your example can be real or imaginary and need not relate to lung cancer or smoking) (**20 marks**).

1. This paper tests the hypothesis that smoking causes lung cancer. Represent this hypothesis more precisely using the counterfactual language and notation as taught in lecture 2. (**10 marks).**
2. Read again the following sections: “Possible causes of the increase” and “Present Investigation”. Now draw a causal diagram using the notation taught in lecture 3.  
   For the purposes of this exercise assume that all variables, when measured, were measured accurately (eg smoking). (**40 marks**).
3. Identify one open pathway through the diagram you have created and explain in plain language how the exposure and outcome might be associated (except via the causal pathway) (**10 marks**).
4. The odds ratio for the association between smoking and lung cancer can be calculated using the data in Table IV of the paper. This gives an odds ratio of 14.6. Now, assume that due to a clerical error cases and controls were mislabelled; 20% of recoded cases were actually controls and 20% of recorded controls were actually cases.
   1. What epidemiological concept does this illustrate (**10 marks**).
   2. How would this clerical error affect your interpretation of the odds ratio for the association between lung cancer and smoking? (**10 marks**).