# Red Meat and Colon Cancer – Exam INTHE4021, H22, Answers

Question 1 counts 15% of the whole exam.

Question 2 counts 15% of the whole exam.

Questions 3-11 counts 70% of the whole exam (i.e. 7.8% each)

## Background

We are interested in the effect of red meat intake on colon cancer among subjects above 40 years.

Based on dietary studies we believe that 20% of the population have a high intake of red meat (>100 Kg/year). The remaining 80% of the population have a normal to low intake of red meat (<100 Kg/year).

Based on register data (Cancer registry of Norway) we think that the overall rate of colon cancer is about 100 cases per 100 000 person years. This puts Norway very high in the world rank for colon cancer. We want to estimate the total effect of red meat intake on colon cancer.

## Design

Research question: “Is high intake of red meat a risk factor for colon cancer?”

1. Given that you have unlimited time and money, please explain how you could get answers to the research question using different designs: cross-sectional, cohort, case-control and ecological design? (maximum 1 page)

Cross-sectional design: draw a random sample of adults, ask questions about intake of red meat (for example last week), and identify participants with colon cancer. Compare the prevalence of colon cancer among those who have high intake of red meat with those who have lower intake of red meat. Will need a very large sample, as colon cancer is a relatively rare disease. The design is prone to information bias – difference in recall of red meat intake among those with cancer and those without cancer? The design cannot infer causality – not possible to know ‘what comes first’ – exposure or outcome? The measured intake of red meat may not reflect the actual exposure when the disease process started (long time ago).

Cohort design: draw a random sample of adults, measure intake of red meat during one week (food-frequency report). Exclude those who already have colon cancer. Follow-up those without colon cancer for a given time period. Compare the risk (incidence) of colon cancer among those with high intake of red meat (exposure) with the incidence of colon cancer among those with lower intake of red meat. Will need to select a very large sample, and follow-up for a very long time. Loss to follow-up will be a problem. Although we measure exposure before the disease occurs, the measured intake of red meat may not reflect the actual exposure when the disease process started (long induction time). Alternatively, link participants from a large health survey conducted long time ago. A health survey, in which intake of red meat was measured. Link data from the health survey with information from the cancer registry or cause of death registry.

Case-control design: Identify cases with colon cancer (for example from patient records in selected hospital(s) or from the cancer registry). Select randomly controls, matched for age and sex, from the population in the catchment area of the hospital(s) / cancer registry. I.e. controls that would have been registered with patient records in the hospital(s) or registered in the cancer registry if they get colon cancer. Ask both cases and controls about consumption of red meat. Estimate the association between high red meat intake (exposure) and colon cancer (outcome). May introduce information bias – difference in recall of red meat intake among cases and controls.

Ecologic design: Correlate amount of red meat sold in different geographical areas (for example municipalities of Norway) the last year(s) with proportion of deaths due to colon cancer in the same geographical areas in the same year(s). Alternatively, measure amount of red meat sold in different geographical areas in different time periods and the proportion of deaths due to colon cancer in the same geographical areas in the same time periods – is there an increase in the proportion of deaths with increasing consumption of red meat? Or a decrease in the proportion of deaths with decreasing consumption of red meat? The design is not suitable for inferring causality, there is no information on individual level, and it is prone to confounding.

1. Which design will you regard as ‘the best’? Discuss why the design is better than the others are? (maximum 1 page)

Practically, a case-control design is ‘the best’ because colon cancer is a relatively rare disease. In general, scientifically, a cohort design is regarded as better than a case-control design, but it has limitations when studying rare diseases – as discussed above (under question 1).

## Measures

Based on a (hypothetical) cohort with 10 year of follow up we have the following data:



Altogether we follow 100 000 subjects, 20 000 have a high(+) intake of red meat, 80 000 have a low(-) intake of red meat. At the end of follow up, a 1000 subjects have developed colon cancer(+) and 99 000 have not(-).

1. Calculate the risk (cumulative incidence), rate (incidence rate) and odds of colon cancer for subjects with: a) high(+) and b) low(-) red meat intake.
2. Use the risk, rate, and odds to calculate the difference and ratio measures using low meat intake as the reference.
3. Interpret the Risk-Ratio, Rate-Ratio and Odds-Ratio in words.

**Answers:**



Risk Ratio: Those who have a low intake of red meat have a risk of 0.6% of colon cancer. Those who have a high intake have 4.7 times the risk.

Rate Ratio: Those who have a low intake of red meat have a rate of 6 cases of colon cancer per 10 000 person years. Those who have a high intake have 4.75 times the rate.

Odds Ratio: The odds of colon cancer is 4.8 times higher among high- compared to low-red meat intake .

Based on a (hypothetical) Case-Control study with 1000 cases and 1003 controls we have the following data:



1. Calculate the (pseudo-)odds for colon cancer for: a) high and b) low red meat intake.
2. Calculate the Odds-Ratio using low meat intake as the reference. Comment on the value compared to the cohort Odds-Ratio (see question 4).

Answers:



The OR from the case-control study (4.38) is close to the OR from the cohort study (4.8). The difference is due to random error in the (somewhat) low number of exposed controls. (Using two controls per case might be better.)

## Causal graph, DAG

The following graph describes the causal relationship between five variables in the analysis.



Age affects both red meat intake and the risk of colon cancer. Red meat intake affects the microbiota (the bacteria composition) in the gut, which in turn affects the risk of colon cancer. And (in a bit of a stretch) both red meat intake and colon cancer affects the economy of the family and the ability to save money. In the following we want to estimate the total effect of red meat intake on colon cancer. The DAG is simplified for use in this exercise, you may comment on missing elements if you wish, but the analyses below should all be based on the present DAG.

1. Is age a confounder, mediator, or collider? Should we adjust for it?
2. Is microbiota a confounder, mediator, or collider? Should we adjust for it?
3. Is saving money a confounder, mediator, or collider? Should we adjust for it?

Answers:

Age is a confounder. We should adjust for it.

Microbiota is a mediator. Since we want the total effect, we should not adjust for it.

Saving money is a collider. We should not adjust for it.

## Regressions

The following shows four logistic regressions based on the Case-Control data.

Model 1:



Model 2:



Model 3:



Model 4:



1. Based on the DAG and that we want to estimate the total effect of red meat intake on colon cancer, what model is the correct one, 1, 2, 3, or 4?

Answers:

Based on the DAG, model 2 (adjusted for age) is the correct model.