'Smokers are more likely than non-smokers to get lung cancer; therefore, smoking causes lung cancer.' Does this statement make sense?

Among the leading evidence that led to the confirmation of the link between smoking and lung cancer in the 1950s was strong **statistical proof** that smoking increases the probability of getting lung cancer through cohort studies. This essay will argue that this statement alone is insufficient for us to conclude that there is a causal relation between smoking and lung cancer, by first showing the **shortcomings of this probability-raising account of causation**, and second pointing out that, even with **other complementary requirements**, including interventionist account via animal experimentation and mechanism account via cellular pathology and chemical analysis, **it is still hard to draw a definite conclusion of causation from correlation.**

1. Argument against probability-raising account:

According to the probability-raising account of causation, "smoking causes lung cancer" means the probability of lung cancer (Lc) given smoking (Sm) is higher than the probability of lung cancer given no smoking, i.e. P(Lc|Sm)>P(Lc|-Sm).

It is true that statistical correlation plays a vital role when scientists search for causation,

- **E.g.** the claim 'smoking causes lung cancer' is backed here by the **widely recognised methodology of population studies** in discovering the lung cancer-cigarette link. In the 1950s, multiple population studies using the method of **case-control epidemiology** concluded that "**smokers of 35 or more cigarettes per day increased their odds of dying from lung cancer by a factor of 40**". In this sense, probability raising theories are highly consistent with how we measure and look for causes.

However, the probability-raising account has two major flaws.

Firstly, it does not rule out probabilistic dependence without causation.

- E.g. One of the **alternative explanations** given by the tobacco manufacturers in the 1950s was that lung cancer and smoking may have a common cause. The fact that P (Lc|Sm)>P(Lc|Sm) does not straightforwardly imply smoking is a cause of lung cancer since the probabilistic correlation might be explained by **another common cause C**, for example, smokers' socioeconomic class.
- **A "screening off" process has to be added to address this problem: if P (Lc|Sm&C) = P (Lc|Sm&-C), and P(Lc|-Sm&C) = P(Lc|-Sm), then C "screens-off" smoking from lung cancer.** Therefore, the probability account may be refined by including an additional clause that the relationship should not be 'screened off' by any plausible common cause, which would enable us to rule out some cases of correlation without causation.

Secondly, the probability raising theory does not give a clear definition of what causes are. Even with screening off in place. It only tells us how we search for causes and why we care about them; it remains dubious whether we can rely on probability raising alone to draw reliable conclusions on causation.

2. Argument against interventionist account:

An interventionist account is also commonly cited as evidence for causation. Such account defines causal relations as relations where changing one variable leads to a change in a second variable, including cases of change of probability or incidence. Using the example of smoking, that is, Sm causes Lc if and only if, were we to manipulate Sm, Lc would also change. For example, the fact that 'tobacco juice' had been shown to cause cancer on laboratory

animals and that the smoke of tobacco could cause tumours when smeared on the hairless skins of rabbits and mice managed to convince many of the causal link.

2.1 Although it is true that scientific intervention had played a vital role in confirming the causal relation in practice, **it is often subject to ethical and legal concerns.** In the case of smoking, most of our evidence for smoking causing lung cancer is observational, as we cannot intervene by forcing people to smoke. It remains questionable whether experimental outcomes on animals could be applied to human.

2.2 Furthermore, intervention must meet stringent requirement to be taken as reliable results. According to Woodward, an intervention I on X must satisfy three conditions: i) the change in the value of X is totally due to the intervention; 2) the intervention will affect the value of Y, if at all, just through the change in the value of X; 3) the intervention is not correlated with other possible causes of Y (other than X). There is often doubts over whether all of the three are satisfied, leaving the result inconclusive. In the case of smoking, such inconclusiveness explains why the tobacco manufacturers in 1953 were able to refute the accumulating evidence and call for 'more research' and successfully convince consumers to go back to smoking.

2.3 Even if we ignore the above limitations, one question remains: **isn't the intervention account's definition of causation still just correlation in disguise**? Notice the use of the phrase 'leads to' in the definition: the definition would be circular if this phrase itself relies on some form of causal link; otherwise, it is still no more than a description of correlation.

3. Argument against mechanism:

Additionally, production accounts via mechanisms also served to distinguish causation. Instead of detecting causal relation through difference-making relations as above, mechanisms focuses on the connection between cause and effect.

For example, the claim that smoking causes cancer was strengthened when pathologists revealed the capacity of cigarette smoke to cause ciliastasis, which could cause cigarette smoke to become trapped in the lungs, causing cancer at locations where the cilia were damaged.

However, even though a plausible mechanism seems to make it more probable that a correlation is causal by revealing the underlying connection, we cannot neglect the obvious fact that **a mechanism itself relies on multiple causal connections**. Just like the case of an intervention account, mechanisms fail to define what causes truly are; it merely pushes the problem to a more basic level – if we zoom in to the causal connections that constitute a mechanism, there is still no way to distinguish them from correlation – how can we justify that a mechanism is anything more than a series of linked correlations?

In conclusion, although there seems to be various accounts that we have relied on in our search for causation, and many of them seem intuitively sensible, providing a philosophically adequate account for causation is actually an exceptionally challenging task, and none of the above accounts provides a satisfaction criterion of definition.