

Supplementary Written Evidence from Professor Marcus Munafò (RRE0094)

The Chair mentioned the potential promise of AI and machine learning approaches, particularly in the context of big data.

This is true, but it cannot be a mindless process otherwise we risk generating misleading or meaningless results. This turns, to some extent, on the question of whether our goal is *prediction* or *causal inference*.

To give an toy example, having yellow fingers may well predict risk of lung cancer (because smoking heavily causes our fingers to turn yellow). But yellow fingers are not a ***cause*** of lung cancer.

AI / machine learning approaches, which are necessarily agnostic as to any underlying causal model, are therefore better suited to research questions about prediction rather than those testing or seeking to identify causal pathways.

The results could be used to ***generate*** hypotheses relevant to causal questions, but these would then still need to be tested in a principled way that incorporated a causal model / theory.

So AI / machine learning is a valuable tool (and one of those that until recently was not available to us), but it is no silver bullet and ***how*** it is used / how robust the results it generates will be will be subject to all the factors we discussed.

1 December 2021