The Perfidiousness of Data

(links at the bottom as hyperlinks have failed)

How can the very same data be used to come to opposite conclusions? Easy: to make things look good, use relative risk reduction, to make things look bad, use absolute risk reduction. Two papers that do exactly this, using the same data, are perhaps the clearest example Dr No has ever come across of this use, or misuse, depending on your viewpoint, of data. The first paper is an Israeli paper that claims to demonstrate the wondrous effectiveness of the Pfizer Biontech vaccine in a real world mass vaccination setting, the second paper is a European paper that uses exactly the same data, lifted number for number from the Israeli paper, to cast serious doubt on the effectiveness of the vaccine. The Israelis conclude that the vaccine has a "high effectiveness", the Europeans conclude that there is a "lack of clear benefit". To add insult to injury, both conclusions are, to put it politely, misleading.

The trail starts with the Israeli study. With a vaccination programme rolling out across the country like a thousand express trains, Israel's faccinatics were naturally keen to prove the vaccine worked, and they had the means. Israel has exceptionally good centralised health data, collected by its integrated health care provider-payer system, making it easy to design a retrospective cohort study to assess vaccine effectiveness. This type of study design sounds like a contradiction in terms, because cohort studies are prospective, in that they move forward over time, from exposure to disease, in contrast to case-control studies, which are retrospective, in that they identify cases of disease, and controls, and then look back in time to determine exposure. But there is a special case cohort study, where you do a prospective study on historical data. Starting today, you go back through the records to identify the exposed, and some controls, and then, again on the historical data, see how the subjects fared over time. A better name for this type of study is a historical cohort study.

This is what the Israelis did. At some time around the beginning of February (the exact timings are not entirely clear, and this has a bearing on things, as we shall see) the researchers analysed the records of all adult members of one of Israel's four health care provider-payers, identifying all those who had received the Pfizer/Biontech vaccine. They then identified matched controls — matched means the controls had similar baseline characteristics, apart from vaccination status — who had *not* received the vaccine, and finally counted various covid outcomes in both groups. It was then a simple matter of comparing the risk of the various covid outcomes in those who had been vaccinated, and those who had not, expressed as a relative risk, and then as vaccine effectiveness, which is simply one minus the relative risk, or, as the authors call it, the risk ratio. The relative risk reduction, or risk ratio, is calculated using division: the risk in the vaccinated group *divided* by the risk in the unvaccinated group. The division has an important consequence: it hides the absolute numbers. A reduction from two per million to one per million has exactly the same relative risk reduction, or risk ratio, as a reduction from twenty thousand per million to ten thousand per million, fifty percent, even though the two circumstances are clearly very different, with the first having one person benefit, the second having ten thousand people benefit.

Despite suffering from a number of handicaps, notably a very short period of follow up, and considerable crossover from the unvaccinated control group to the vaccinated group, the results seem impressive. Among the swathe of reported vaccine effectiveness, we have, for example, 84% effectiveness against death during

week 4 after vaccination, and 94% effectiveness against PCR confirmed symptomatic illness from 7 days after vaccination to the end of the study. But these *relative* effectiveness figures mask the *absolute* numbers. Despite almost 1.2 million subjects, half vaccinated and half unvaccinated, there were only 41 deaths, 9 in the vaccinated group and 32 in the unvaccinated group. Yes, there is an effect, but you have to vaccinate almost 600,000 people to prevent 23 deaths.

Which is how the European study managed to spin the results the other way. They lifted Table 2 from the Israeli study, deleted the effectiveness column, and added a NNTV (number needed to vaccinate) column, and, at a stroke, or rather two, the same results now look very different. The 84% effectiveness against death in week 4 after vaccination translates into a NNTV of 16,667 (95% CI 9,000 to 50,000, because the numbers are so small). You have to vaccinate tens of thousands of people, with the attendant risks of serious side effects and even possible death, to prevent one death in the fourth week after vaccination. For the 41 total deaths, 9 in the vaccinated group and 32 in the vaccinated group, Dr No estimates the absolute risk reduction to be 0.000038 (32/596,618 minus 8/596,618), giving a NNTV of 25,940.

A vaccine effectiveness of 70% or better (depending on follow up interval) against death, or an overall NNTV of almost 26,000 to prevent one death during the study period? Both are true, but which is right? One supposes it depend on whether you are bowling from the Nursery End or the Pavilion End. Or perhaps both are wrong, not in the numbers (which remain correct wherever one bowls from), but in the message. Dr No has on a number of occasions dropped in hints to a major flaw that compromises both studies: the ridiculously short follow-up period. The mean follow-up period was a tiny 15 days (interquartile range 5 to 25 days), a ludicrously small period for a vaccine. This meant there were very few deaths (41 among almost 1.2 million subjects) but the Israeli study masks this by using relative numbers. If there had been just 14 deaths, three in the vaccinated group and 11 in the unvaccinated group, they would have reported very similar results, because the relative differences, 9 to 32 and 3 to 11, remain the same. We can almost say the same thing about a total of five deaths, one in the vaccinated group and four in the unvaccinated group. The message is clear: *never* consider a relative risk reduction, or a derived vaccine effectiveness, without knowing the context, numbers treated, number of outcomes.

The very same flaw in the Israeli study, inadequate follow-up, fires in exactly the opposite direction in the European study. Because the follow-up period was so short, very few deaths were likely to occur, whether the subject was vaccinated or not, and this smoothly translates into huge NNTVs, 9,000 or 16,667 or 25,940 or 50,000, take your pick, it doesn't really matter, as all of them make the point that huge numbers need to vaccinated, and so put at risk of vaccine related harm, to save one life *in the short time frame of the study*. NNTs (and so NNTVs), because they are based on incidence (the number of new cases in a defined population *over a defined time period*) are *always tied to a time frame*. For a vaccine, which if it works can be expected to continue to work over time, and so produce more beneficial outcomes as time goes by, a mean follow up of 15 days is comically short. So just as we need to be careful with relative risk based estimates (know the context), so too with NNTs do we need to be careful (know the time frame).

Bottom line? Bin both papers.

Israeli Study: https://www.nejm.org/doi/10.1056/NEJMoa2101765
European Study: https://www.mdpi.com/2076-393X/9/7/693/htm