Charlie Rettig, Johannes von Kleist, Marcus Wenau

Intro to Focus Areas CS Exercise 4

Task 1

a)

The objective of Task 1(a) was to model drug uptake and elimination in the body using an ODE solver. The reaction rates used for this model are:

$$r_0 = x_0 * k_a, \quad r_1 = x_1 * k_e$$

Where k_a = 0.5, k_e = 0.3, initial dosage $x_0(t_0)$ = 200 mg, and initial bloodstream concentration $x_1(t_0)$ = 0 mg/L.

x_0 is the amount of the drug consumed while x_1 is the concentration within the bloodstream. Therefore r_0 represents the rate at which the drug enters the bloodstream, while r_1 is the rate at which the body absorbs the drug from the bloodstream.

To verify our result we were supposed to plot our results compared to the analytical solution. The analytical solution for x_1(t) is given by:

$$x_1(t) = x_0(t_0) * \frac{k_a}{(k_a - k_e)} * (e^{-k_e * t} * e^{-k_a * t})$$

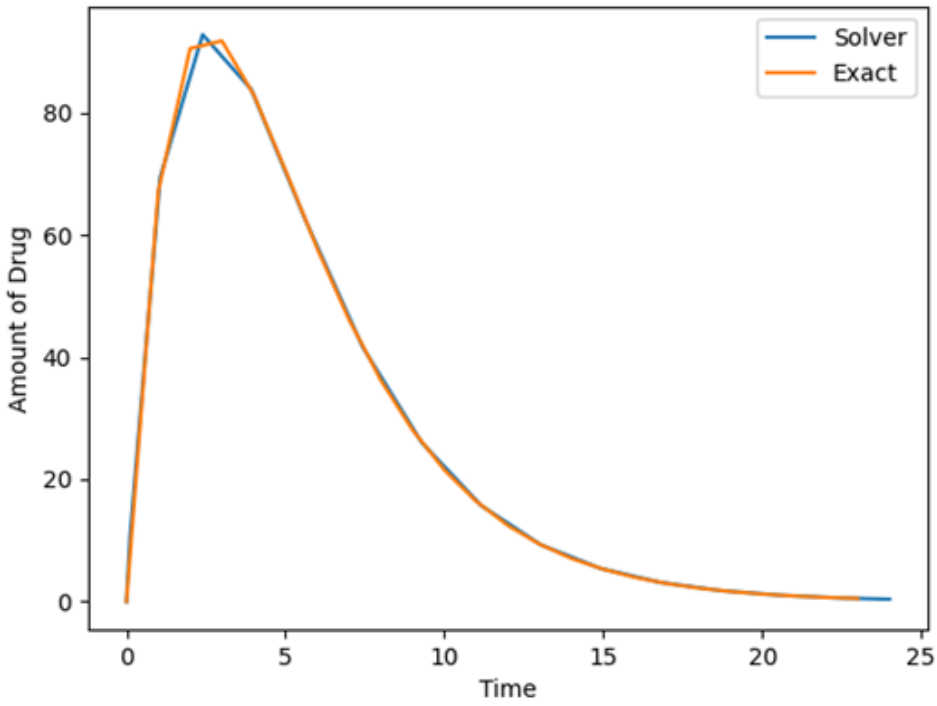


Figure 1: Comparison of analytical and numerical solutions for the ODE

The comparison can be seen in the plot above in Figure 1. The analytical and numerical solutions look quite similar.

b)

In Task 1 b) we altered the program from Task 1 a) so that it simulates repeated infusions every 12 hours for 7 days. In doing so we were supposed to answer 4 questions:

The first was to find out at what point the pharmacokinetic steady state is achieved. This can be found out by comparing the value x_0 after receiving a dosage; if this value is equal for 2 doses in a row, the pharmacokinetic steady state is reached and the values will continue to stay the same. This happens at the 11th dosage, in other words after 120 hours.

The second question asked for the minimum (trough) concentration level reached after the pharmacokinetic steady state is achieved. To find this value one must simply check the lowest value x_1 reaches after the 11th dose which happens to be 12.80 mg/L

Similarly to question 2, questions 3 and 4 asked for the maximum and mean values after the pharmacokinetic steady state was reached.

The maximum value of x_1 or in other words the maximum concentration in the blood reached was 98.50 mg/L while the mean after the 11th dose was 46.53 mg/L

We also plotted the graph for x_1 with the parameters given in Task 1 b) as can be seen below in Figure 2.

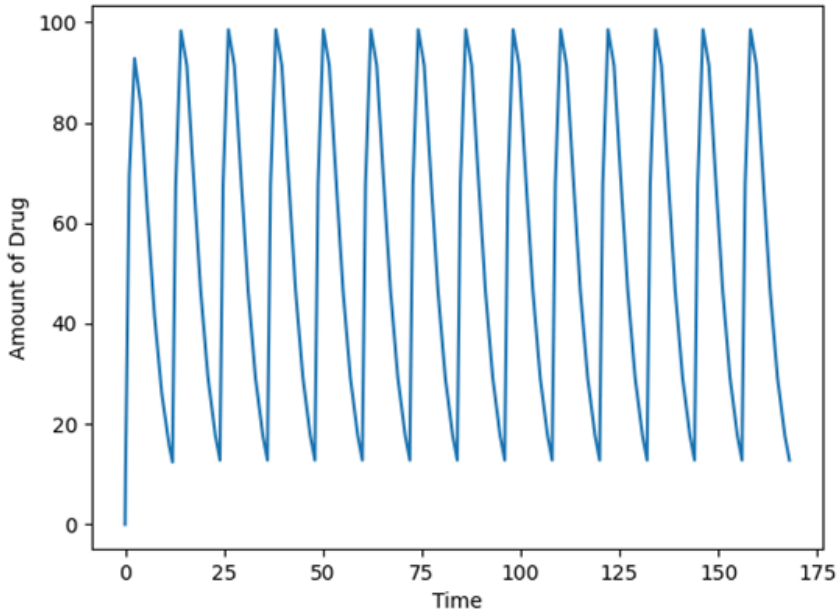


Figure 2: Graph of x1(t) with repeated dosages

Task 2

Table 1: Table of data (number of participants, average follow-up time, and number of infections) for each vaccine trial, differentiated into treatment and placebo arms

Trial	Number of participants (N)	Average follow-up time per person in years ($\widehat{T}$)	Number of Infections (I)
Astra-Placebo	8550	0.1778082191780822	130
Astra-Vaccine	17662	0.177534246575342	73
Moderna-Placebo	14073	0.3383117989698468 7	269
Moderna-Vaccine	14134	0.407356428298837	19
BionTech-Placebo	17511	0.1268916680943407	162
BionTech-Vaccine	17411	0.127160990178623	8

a)

Task description: Using the above data in Table 1, the goal was to inversely sample one thousand incidence rates for the placebo arm of each trial. The provided program IncidenceParameter.py was used to calculate the incidence rates for each trial.

Results and discussion:

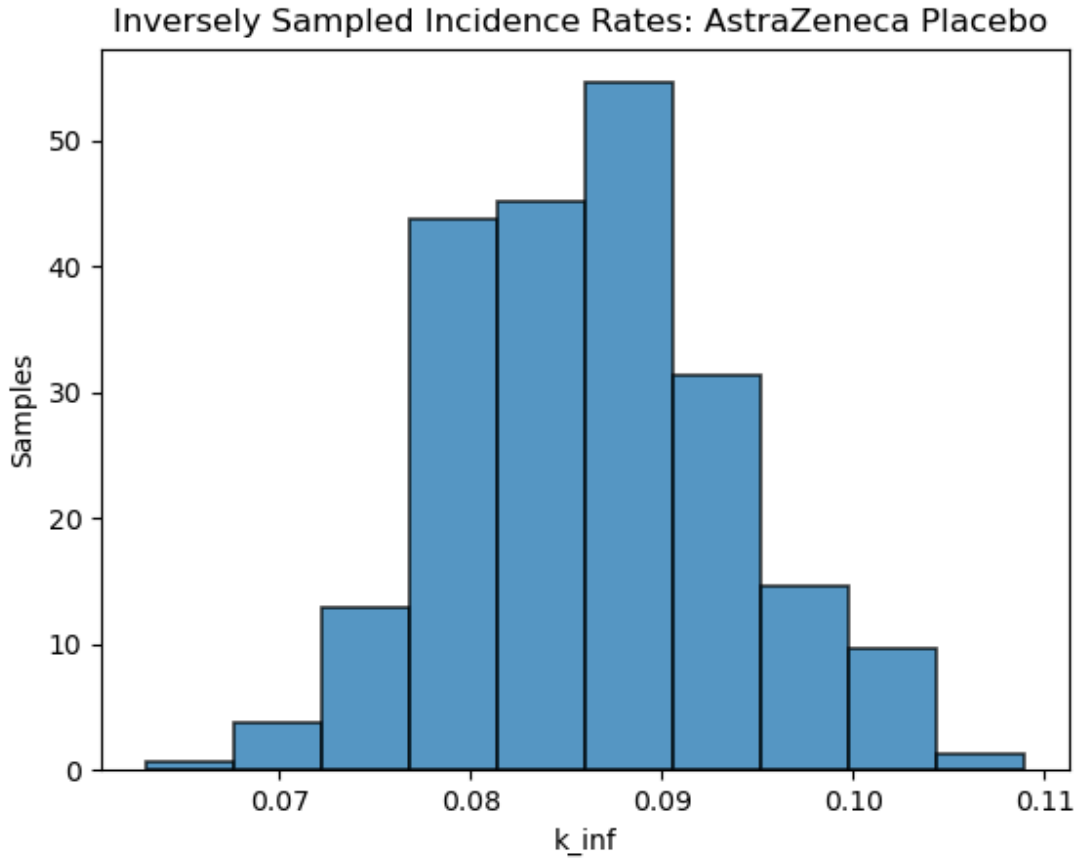


Figure 3: Histogram of the incidence rates for the placebo arm of the AstraZeneca trial

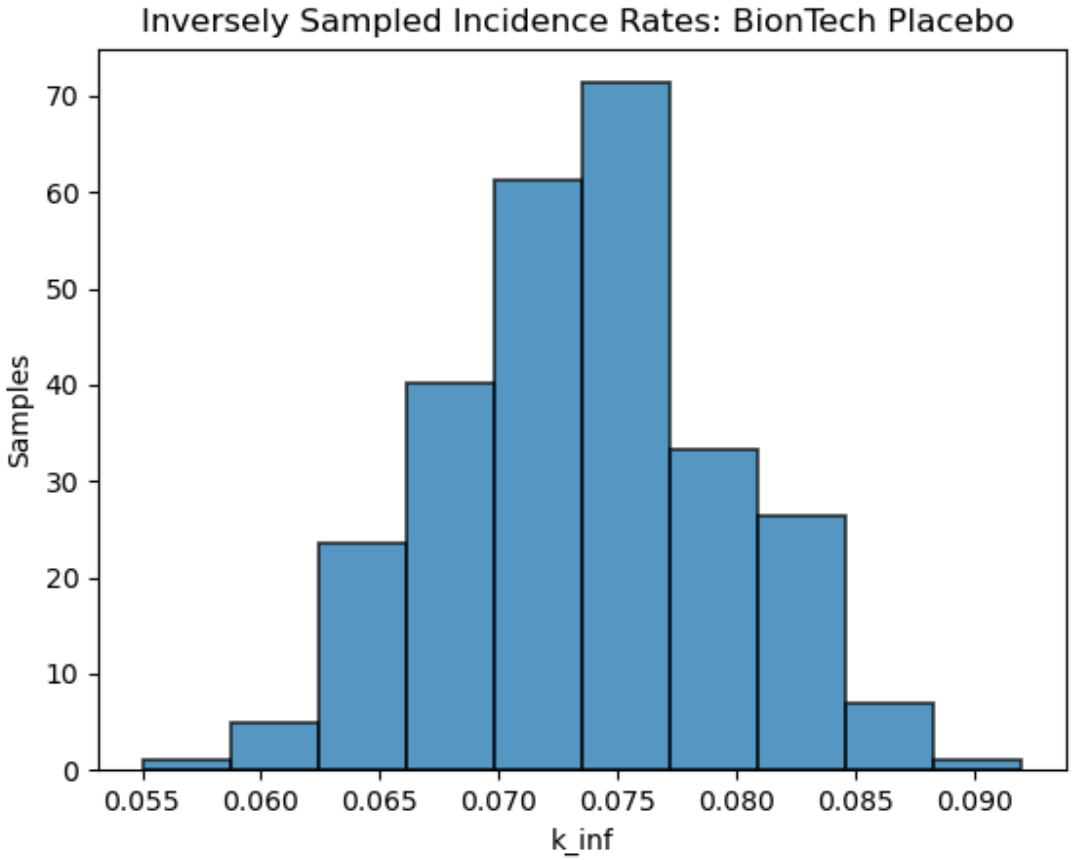


Figure 4: Histogram of the incidence rates for the placebo arm of the BionTech trial

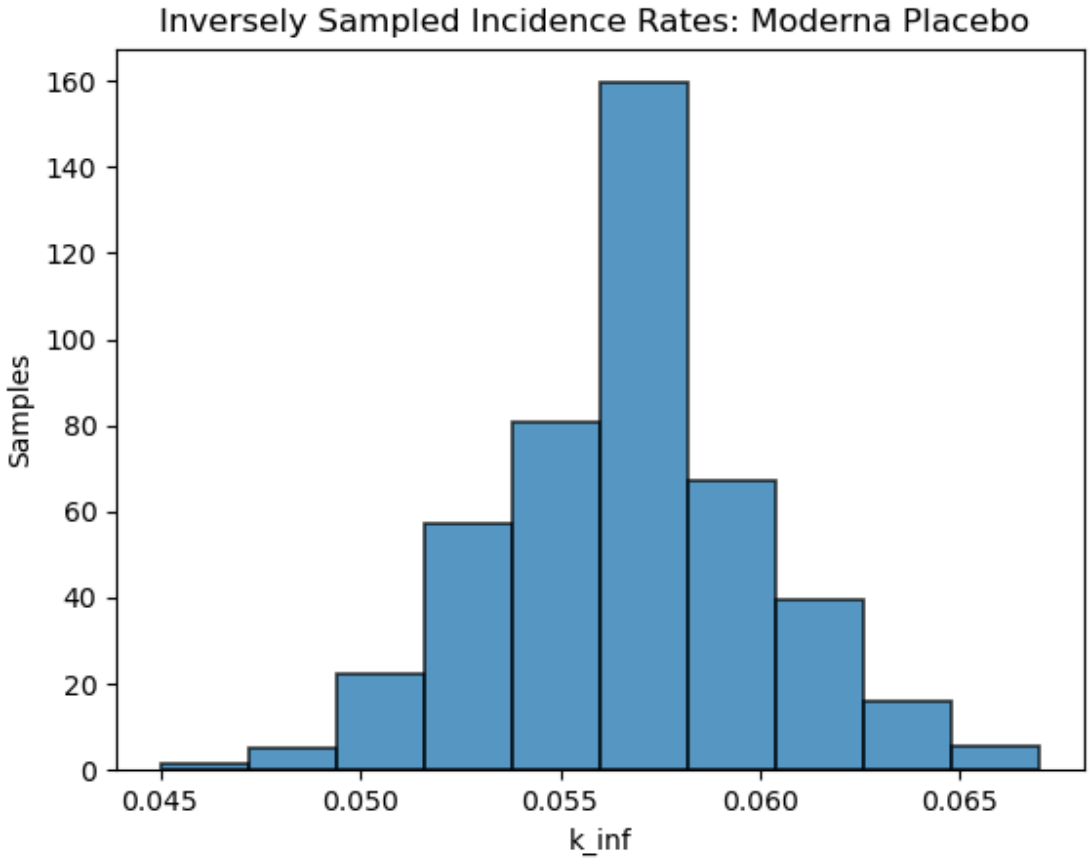


Figure 5: Histogram of the incidence rates for the placebo arm of the Moderna trial

In the three figures 3-5, the respective sampled incidence rates for the placebo arm are plotted in histograms. All the results across all histograms are within one order of magnitude (between 0.04 to 0.11) with an approximately normal distribution of the results, with the visual discrepancies from the normal distribution being in part due to arbitrary bucketing in the histograms and of course the stochastic nature of the sampling. The differences between the placebo arms of the different trials can be explained by a variety of factors: the populations are of course different between the trials, different geographic locations which might have different susceptibilities and base incidence rates as well as differences between individuals, one would hope no individual took part in more than one Sars-Cov-2 vaccine study. Additionally, these geographic differences might also imply different strains, with different levels of infectiousness, were present during the trials. These differences in infectiousness could also be caused by temporal differences: AstraZeneca from August 2020 to January 2021, whilst Moderna and BioNTech had their first injections take place from July to October and November, which might also explain why the AstraZeneca trial had a larger difference in incidence rates in comparison to the other two. The histograms for the treatment arms of the trial are listed further below.

b)

Task description: The placebo arm of the three vaccine trials of interest were simulated 100 times using stochastic simulation. Incidence rates were sampled in task 2a), from which drop-out rates were calculated using the following formula: 'average follow-up time per person' = 1/(kdr−out+kinf). Average follow-up time per person was extracted directly from the respective literature for each study.

Discussion: The histograms shown below represent the distribution of incidence rates at the end of each simulation. Summary statistics are displayed below each plot along with bootstrapped 95% confidence intervals.

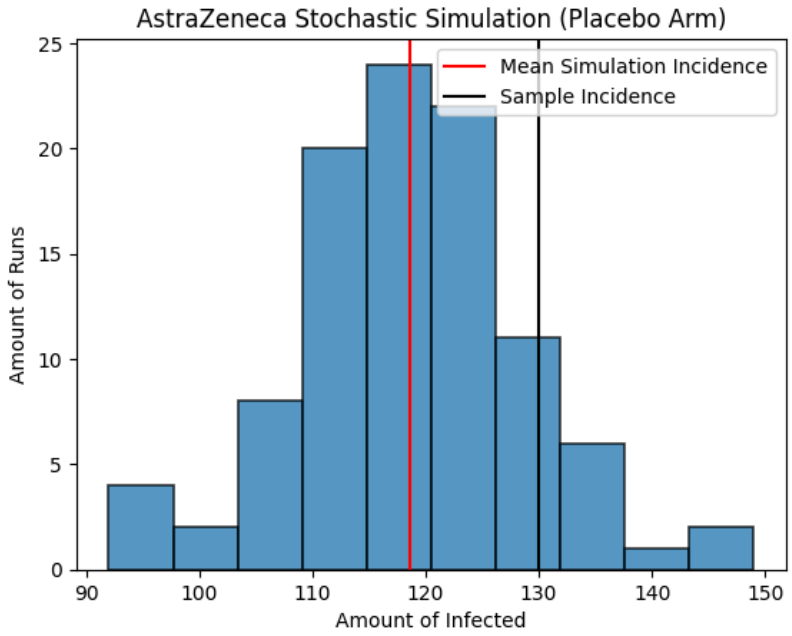


Figure 6
Mean: 118.6349 (116.7, 120.6)
Bootstrap Standard Devation: 10.0097242137783 (8.4521, 11.6216)

For the placebo arm of the AstraZeneca trial, the mean of the simulation incidences was an underestimate of the sample incidence rate (130) as the sample rate is not captured in the bootstrapped 95% CI for the simulation mean. There is low uncertainty (standard deviation ≈ 10.01) in the predictions, indicating that there is some underlying issue with the models used. However, this could also indicate that the original study design was also flawed, yielding a misleading sample incidence that does not reflect larger, more diverse populations.

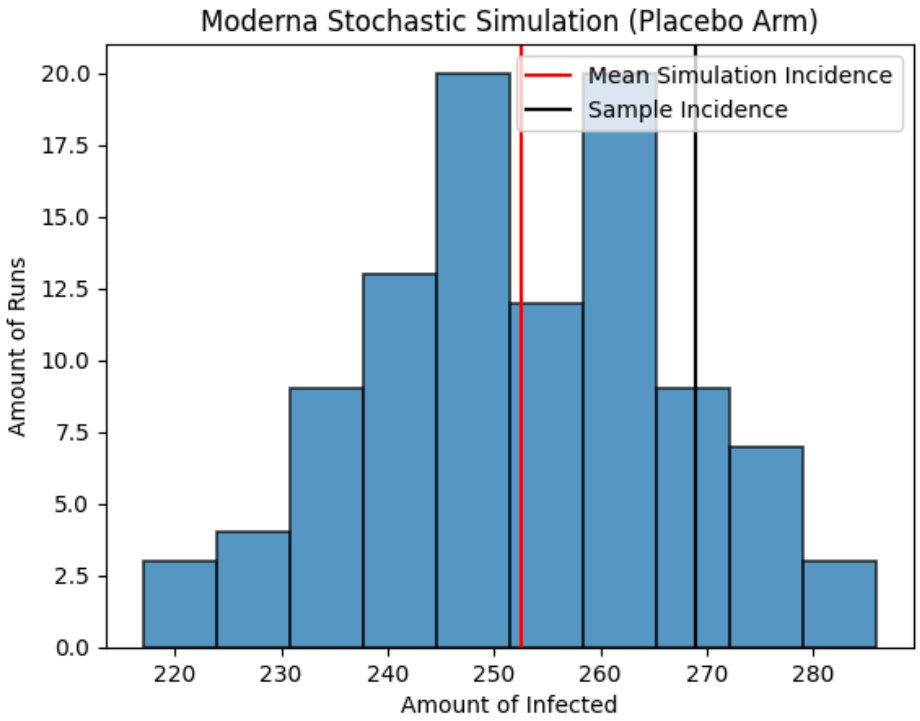


Figure 7
Mean: 254.98 (252.1, 257.9)
Bootstrap Standard Devation: 14.6302 (12.8319, 16.3962)

The placebo arm of the Moderna trial yielded similar results to the AstraZeneca trial in terms of the relationship between the mean simulated incidence rate and the study sample incidence rate (269). The uncertainty of the mean simulated incidence rate, again represented by the bootstrap standard deviation, is 14.63. This is quite low in the context of the range of the data and thus reflects a relatively low uncertainty in the simulation's predictions.

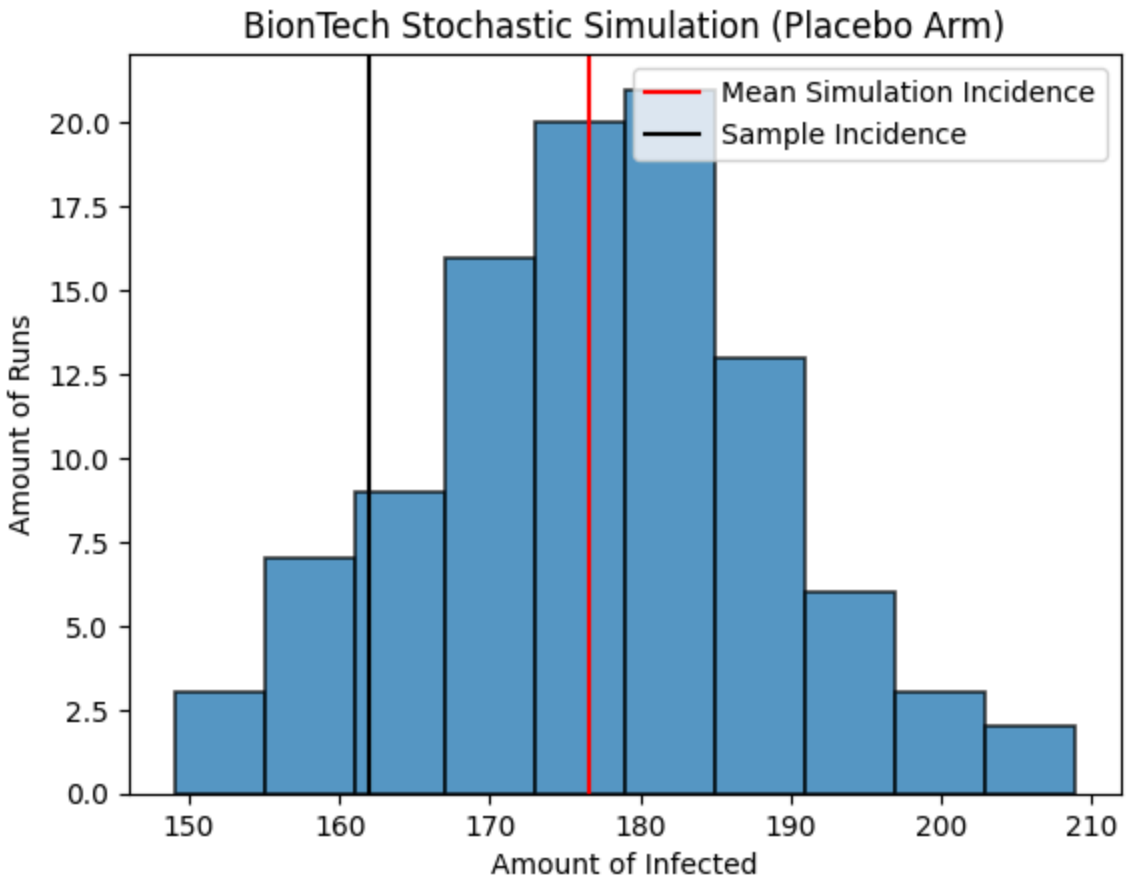


Figure 8
Mean: 176.59 (174.25, 178.99025)
Bootstrap Standard Devation: 11.975618959241647 (10.3841, 13.6016)

In the placebo arm of the BionTech trial, the same result is yielded as the aforementioned studies; the sample incidence (162) is not captured in the 95% CI. Here, the simulation prediction is an overestimate of the sample incidence. The uncertainty (standard deviation) is 11.97, indicating low uncertainty in the simulations' predictions.

Differences between simulation incidence rates and those of their respective studies could arise because of a variety of factors, including, but not limited to:

- Individual effects of Placebo Effect: Some individuals might be more susceptible to the effects of the Placebo effect. The opposite assumption is made in our models.
- Stochastic nature of simulation: The stochasticity in our simulations performed may not accurately reflect real-world stochasticity, as many factors are either over-simplified or ignored in our simulations.
- Individual risk homogeneity: Our models assume uniform risk across all individuals. In real-world scenarios, individual risk is highly variable, especially across different demographics.
- Input data: the data from each study represents a small proportion of all individuals world-wide who are susceptible to COVID-19. Thus, the susceptibilities of those

represented in the three trials might not accurately represent those of individuals outside of the study population.

- Study design: Each study is designed differently. Because each simulation makes identical simplifications and assumptions, this leads to varying discrepancies between the predicted and actual incidence rates.

c) Task description: Using the above data in Table 1, the goal was to inversely sample one thousand incidence rates for the vaccination arm of each trial. The provided program IncidenceParameter.py was used to calculate the incidence rates for each trial. Additionally, the vaccine efficacy should be estimated as well as the uncertainty in that estimate.

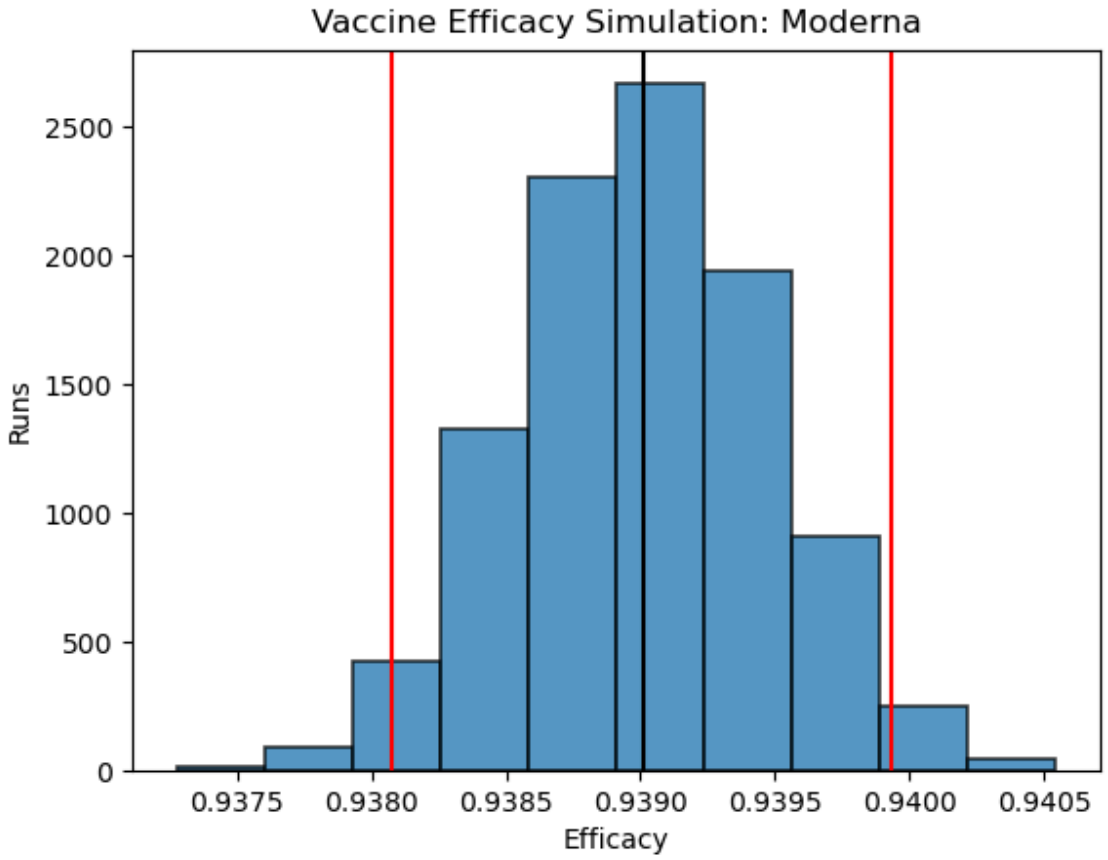


Figure 9: Histogram of the results of the bootstrap-simulation for the efficacy of the Moderna vaccination, mean in black and uncertainty of one standard deviation in red

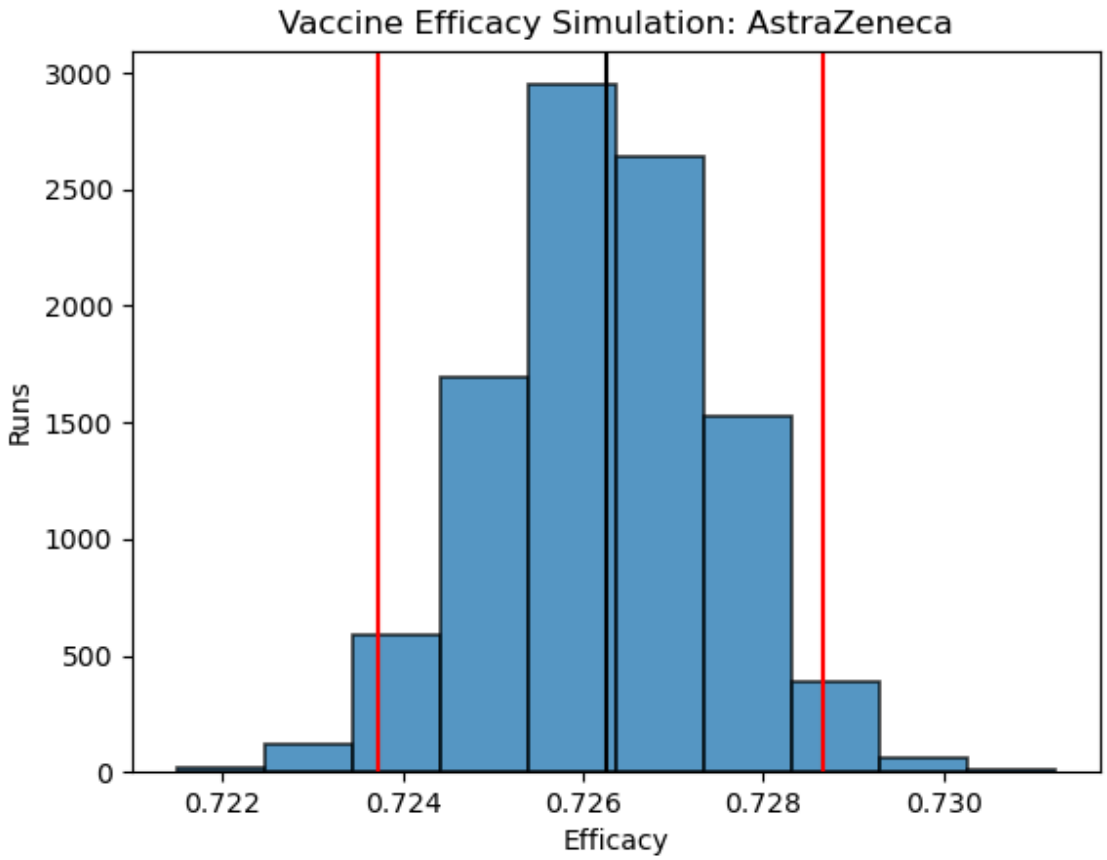


Figure 10: Histogram of the results of the bootstrap-simulation for the efficacy of the AstraZeneca vaccination, mean in black and uncertainty of one standard deviation in red

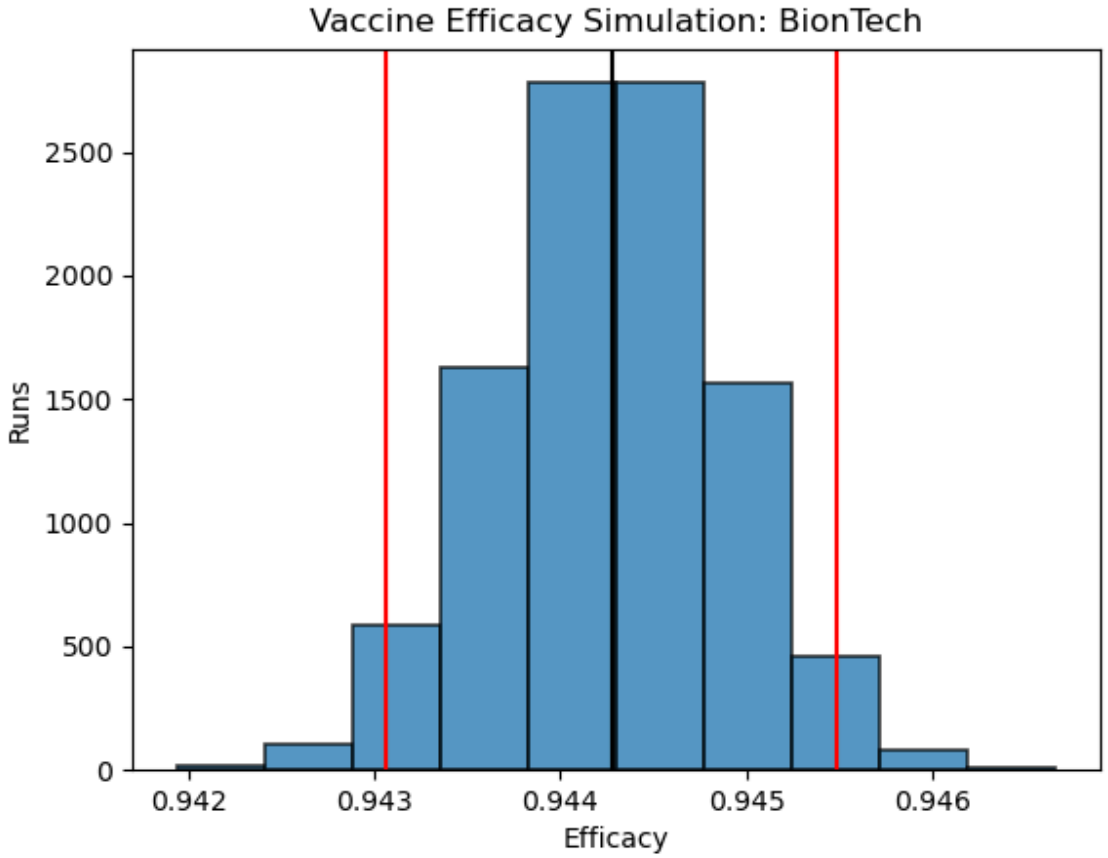


Figure 11: Histogram of the results of the bootstrap-simulation for the efficacy of the BionNTech vaccination, mean in black and uncertainty of one standard deviation in red

Results & Discussion:

Vaccine efficacy was calculated using the following equation: $\varphi = 1 - \frac{k(vacc.)}{k(plac.)}$ with k(vacc.) and k(plac.) representing an incidence rate for the vaccination and placebo arm respectively. The mean of this quotient across all sampled incidence rates was calculated for each trial, resulting in a vaccine efficacy of 72.62% for AstraZeneca, a vaccine efficacy of 93.89% for Moderna, and a vaccine efficacy of 94.43% for BioNTech.

To estimate the uncertainty in this estimate, we employed the method known as bootstrapping, (known from our course Foundations in Mathematics). To this end, we resampled from our incidence rates 10000 times drawing a thousand random samples (with replacement) each. For each of the 10000 runs we then calculated the mean efficacy and from this, calculated an overall mean, an overall standard deviation, as well as a 95% percentile confidence interval. The results of these resampling runs are visible in the above figures, with one histogram per trial. This resulted in very similar means, differing at most by 0.01%, which are included in the histogram as a vertical black line. The boundaries for one standard deviation are marked in red, and are 0.00125, 0.00047, and 0.00061 respectively, for AstraZeneca, Moderna and BioNTech. The 95% confidence interval was excluded from the histograms to avoid visual clutter but are

72.37% to 72.86%, 93.80% to 93.99%, and 94.30% to 94.54% respectively, for AstraZeneca, Moderna and BioNTech. The results match those from the paper relatively closely, with the confidence intervals of the papers being broader than ours, likely due to not being calculated in the same way as ours. The minor differences (less than 1% for Moderna and BioNTech, and less than 2% for AstraZeneca) in the means between simulations and efficacy reported in papers is likely due to the stochastic nature of the simulation and the underlying processes. The fact that they match so closely is also due to the sampled incidences being obtained from the studies themselves directly and as such it is unsurprising, especially with many repetitions of bootstrapping to reduce the variance that the results of the simulation converge relatively closely to the real results.

d)
Task description: The placebo arm of the three vaccine trials of interest were simulated 100 times using stochastic simulation. Incidence rates were sampled in task 2a), from which drop-out rates were calculated using the following formula: 'average follow-up time per person' = 1/(kdr−out+kinf). Average follow-up time per person was extracted directly from the respective literature for each study.

Discussion: The histograms shown below represent the distribution of incidence rates at the end of each simulation. Summary statistics are displayed below each plot along with bootstrapped 95% confidence intervals.

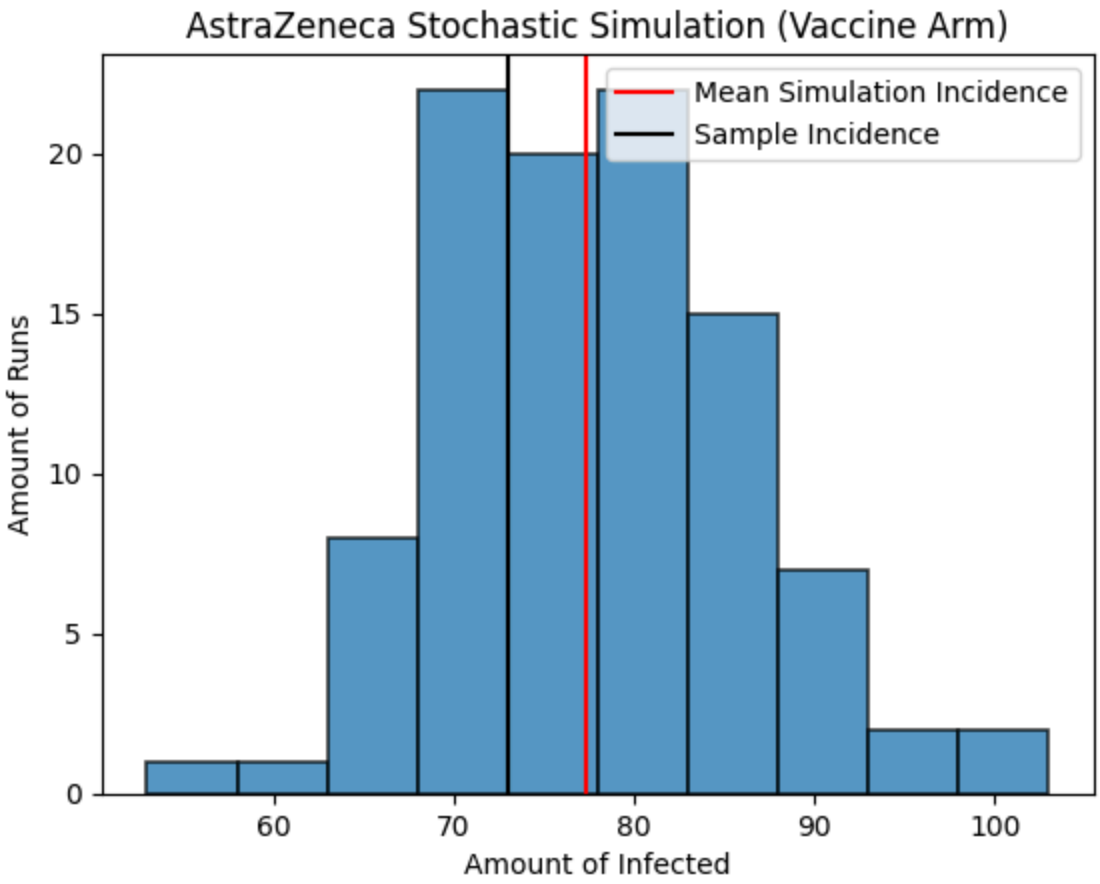
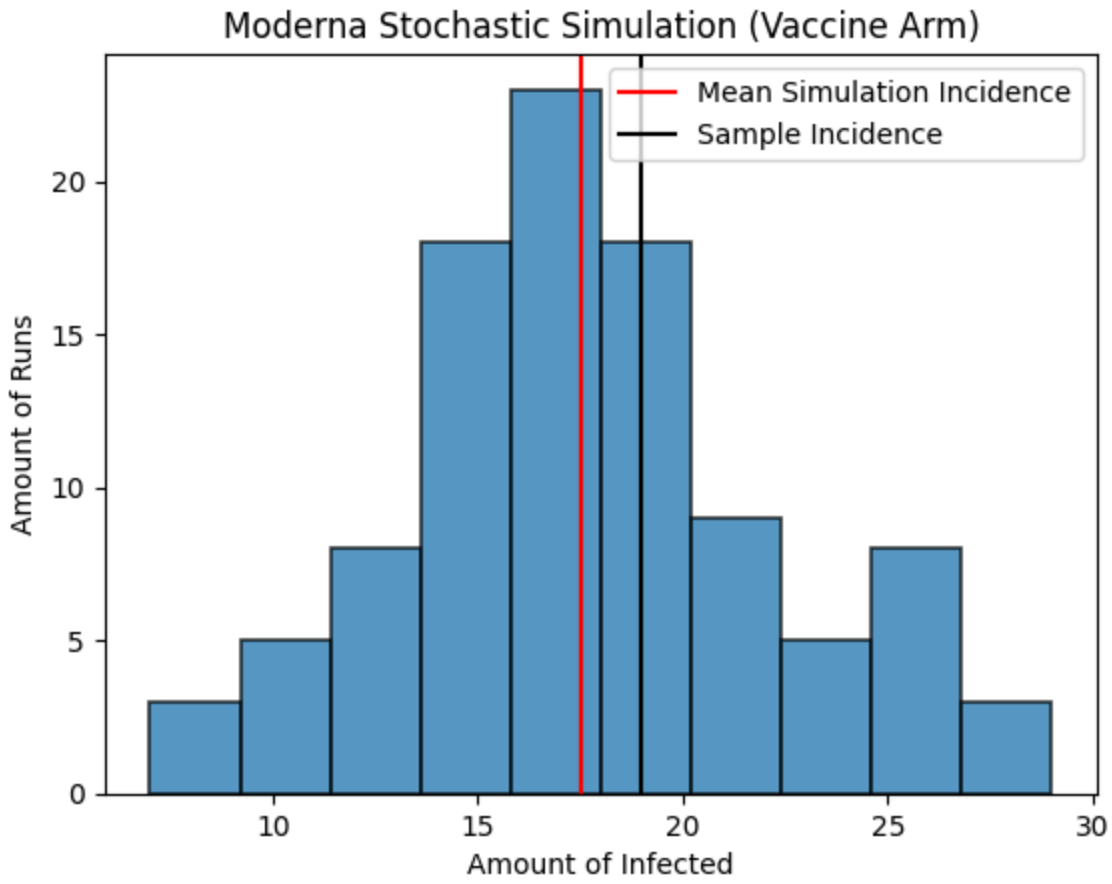


Figure 12
Mean: 77.3585 (75.68, 79.0602)
Bootstrap Standard Deviation: 8.4427 (7.1886, 9.7414)

For the vaccine arm of the AstraZeneca trial, the mean simulation incidence is a slight overestimation of the sample incidence reported in the literature. The sample incidence is not captured in the 95% CI for the mean, the raw predictions come with a relatively low uncertainty (8.4427). Therefore, there is a low probability that any of the factors listed at the end of this section may have reduced the accuracy of the simulations. In any case, if we were to conduct more simulations, the 95% CI for the mean might eventually capture the sample incidence.



Mean: 17.52 (16.65, 18.4202)
Bootstrap Standard Deviation: 4.557 (3.9621, 5.1437)

The simulations of the Moderna trial were a slight underestimate of the sample incidence reported in the trial (19). The sample incidence falls just one incidence outside of the 95% confidence interval for the mean of the simulation incidence rates, suggesting that more simulations might yield an overlap. However, this might be less likely given the low uncertainty of just 4.557.

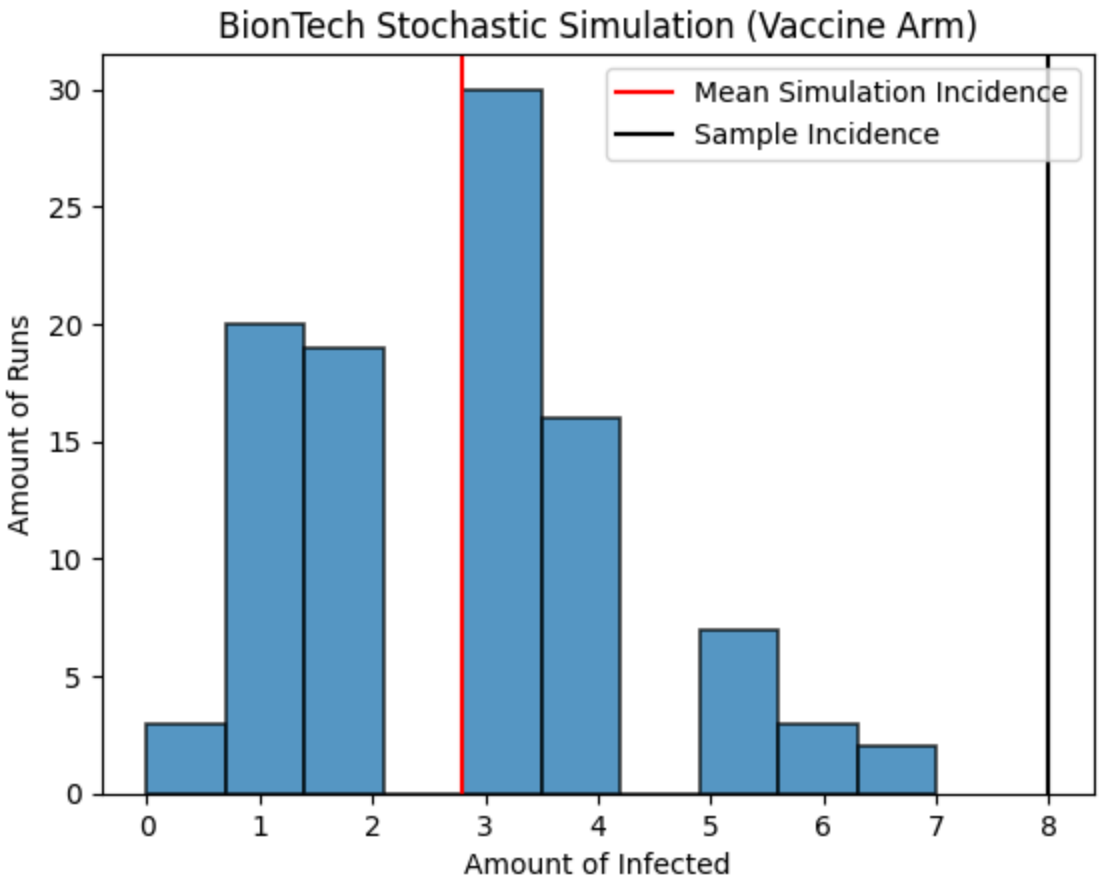


Figure 13
Mean: 2.79 (2.5, 3.08)
Bootstrap Standard Deviation: 1.4964 (1.2874, 1.7078)

For the simulation of the vaccine arm of the BionTech trial, the simulations seemed to severely underestimate the incidence rate; the sample incidence rate is more than 2 standard deviations away from the mean and is not captured in the 95% CI of the mean.

Differences between simulation incidence rates and those of their respective studies could arise because of a variety of factors, including, but not limited to:

- True Vaccine Efficacy: The simulated vaccine efficacy might differ from the real-world efficacy due to variations in factors such as differing immune response among different populations and varying effectiveness against specific variants.
- Stochastic nature of simulation: The stochasticity in our simulations performed may not accurately reflect real-world stochasticity, as many factors are either over-simplified or ignored in our simulations.
- Individual risk homogeneity: Our models assume uniform risk across all individuals. In real-world scenarios, individual risk is highly variable, especially across different demographics.
- Input data: the data from each study represents a small proportion of all individuals world-wide who are susceptible to COVID-19. Thus, the susceptibilities of those

represented in the three trials might not accurately represent those of individuals outside of the study population.

- Study design: Each study is designed differently. Because each simulation makes identical simplifications and assumptions, this leads to varying discrepancies between the predicted and actual incidence rates.

Appendix

Listed here are the figures not included in the immediate report due to not being relevant for the questions, but still relevant for complete context (for example due to the data being used later on)

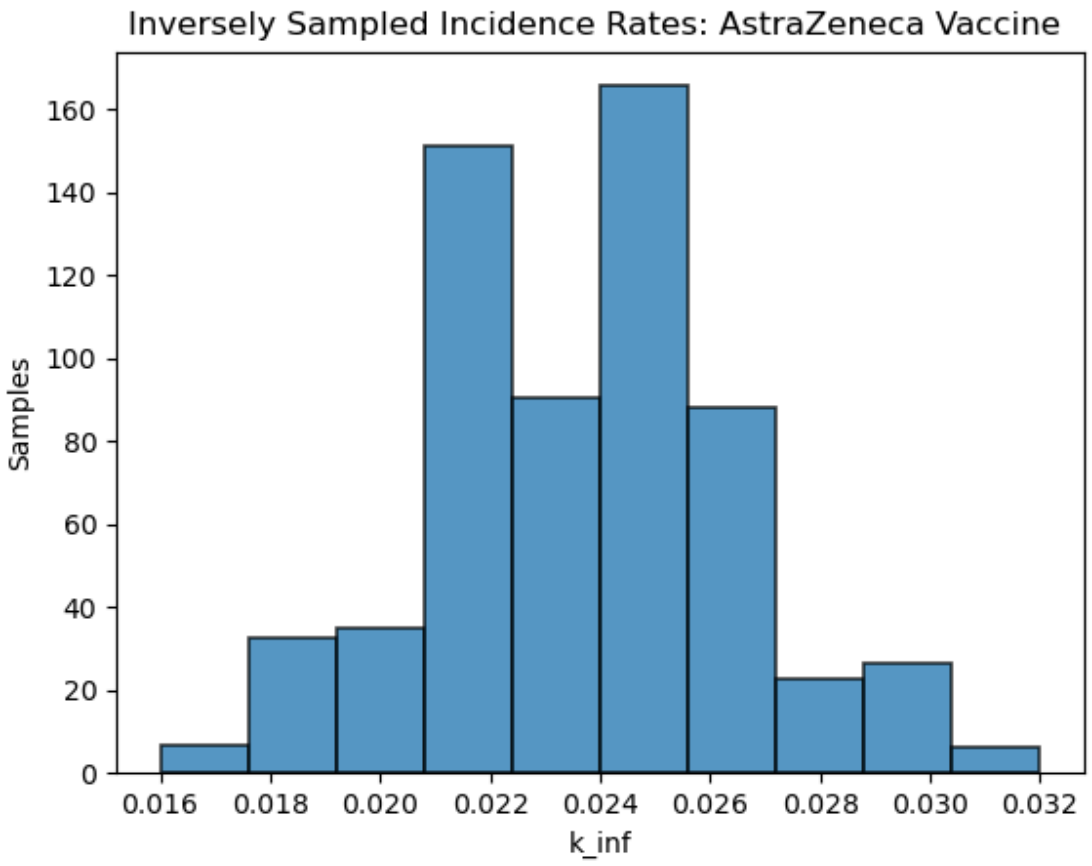


Figure 14: Histogram of the incidence rates for the vaccination arm of the AstraZeneca trial, data used in d)

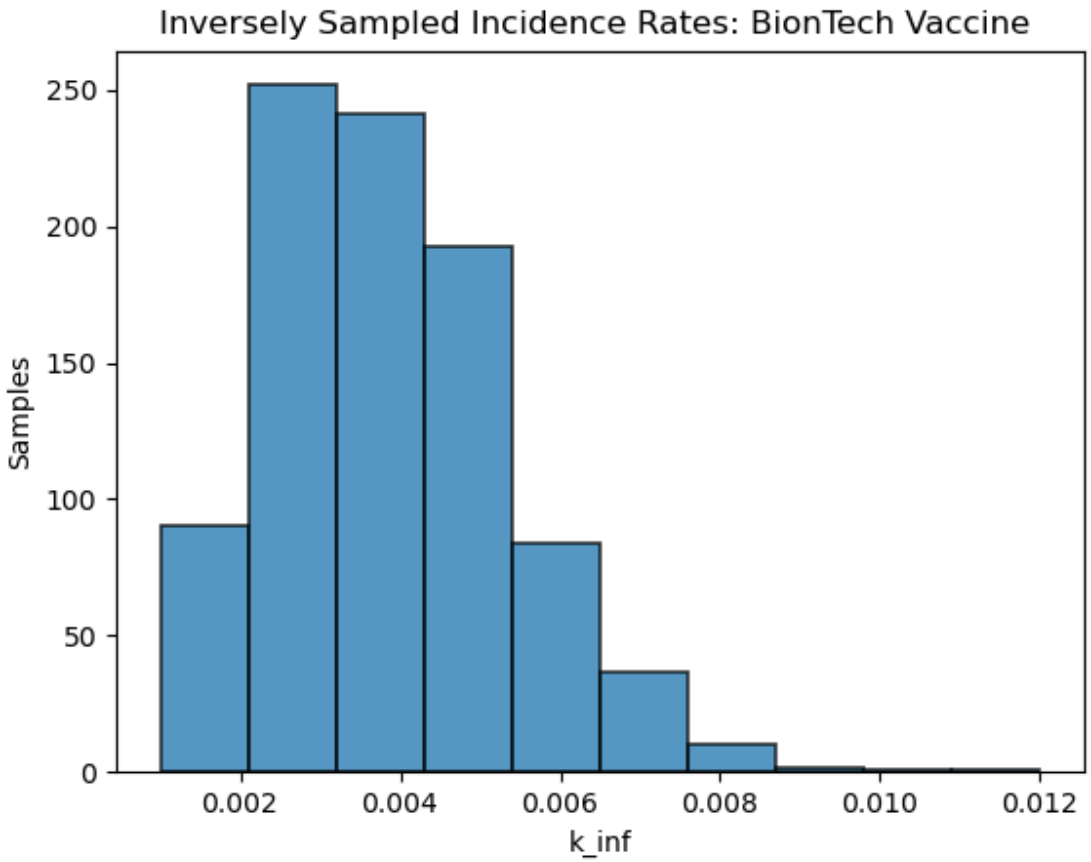


Figure 15: Histogram of the incidence rates for the vaccination arm of the BionTech trial, data used in d)

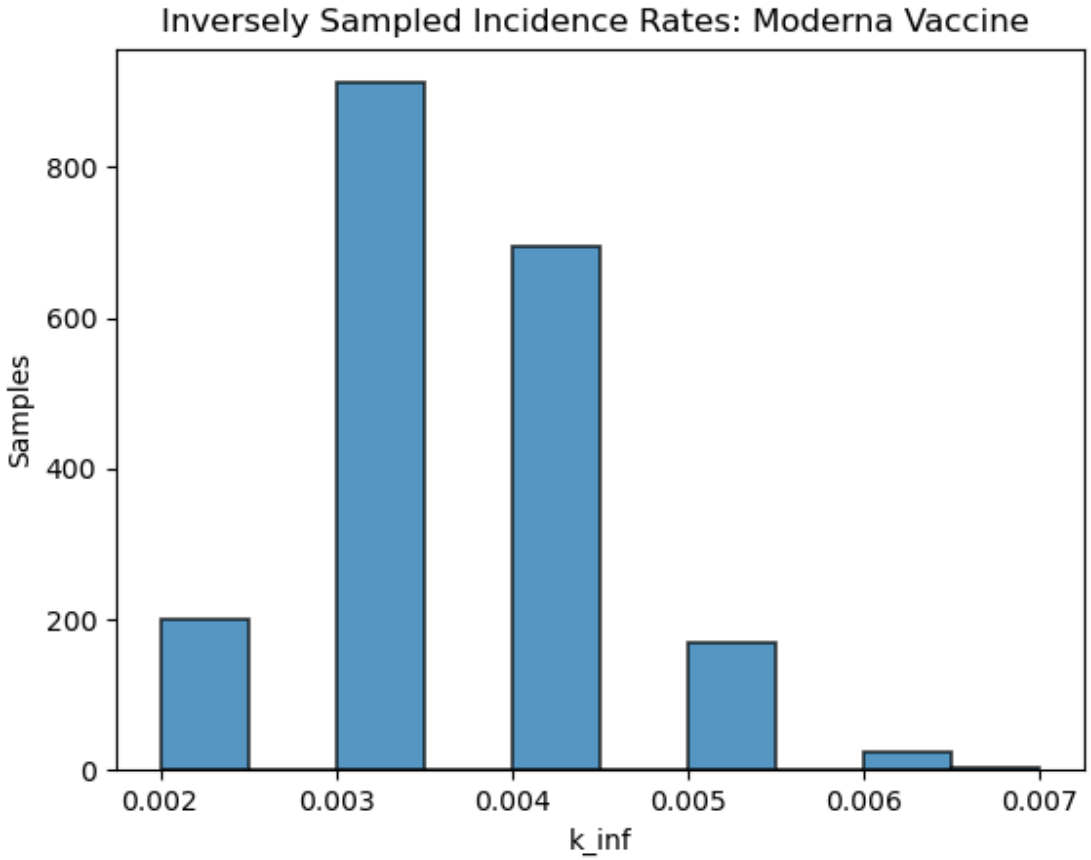


Figure 16: Histogram of the incidence rates for the vaccination arm of the Moderna trial, data used in d)