

Statistical Approaches In Complex Efficacy Analysis for Vaccine Trials

Nijitha L M
Drisya C A
Genpro Research

Vaccine Efficacy



Vaccine Efficacy can be estimated as one minus some measure of relative risk(RR) in the vaccinated group compared to the unvaccinated group .

$$VE=1-RR$$

Vaccination protects vaccinated individual !

- VE_S - Vaccine Efficacy for Susceptibility

$$VE_S = 1 - \frac{R(\text{Vaccinated People})}{R(\text{Unvaccinated people})}$$

Where R is a measure of risk

- It can be Hazard ratio, Incidence rate or cumulative incidence

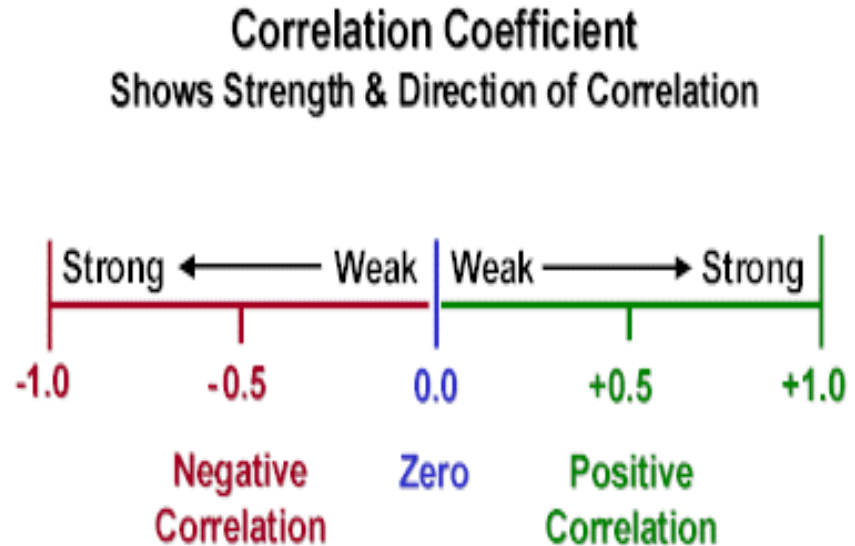
- VE_I = Vaccine Efficacy for Infectiousness, by comparing per –contact transmission probability from vaccinated people who become infected with transmission probability from unvaccinated people who become infected
- VE_P = Vaccine Efficacy for Progression, effect of vaccine in reducing the severity of disease or probability of death in symptomatic cases

Influencing Factors

- **Demographic Factors** – Age, Sex, Race, Ethnicity , BMI etc
- **Behavioural Factors**- Smoking, Stress, Depression etc
- **Environmental Factors**- Geographical Region, Season, Herd immunity etc
- **Other factors**- Vaccine Type , co-administered drugs, genomic mutations of virus /new variants

Correlation

- Association between two quantitative variables.
Assume that the association is linear, that one variable increases or decreases a fixed amount for a unit increase or decrease in the other.
- Scatter Plot Method
- Correlation Coefficient



Significant Covariates Identification

- Regression Analysis

Fitting the regression line

- $Y = \beta_0 + \beta_1 x + \varepsilon$
- we assume that $\varepsilon \sim \text{i.i.d}$
 $N(0, \sigma^2)$
- Based on the data and distribution of variables we can select an appropriate model to fit the data.

Model Validation Techniques

- R^2 - This is the value that is
- Used for predicting the extent of variability in the dependent variable that can be explained by the independent variable
- Residual Plots
- AIC, BIC

Simulation Study

- Simulated a data of 1000 observations based on covid -19 data.
- Our variable of interest is occurrence of disease in vaccinated and unvaccinated group. Since the dependent variable is a binary one, we used logistic regression model.
- Used the below equation to find out the vaccine efficacy.

$$IRR = \frac{OR}{(1 - P_i) + (P_i * OR)}$$

where p_i = overall incidence rate in population

- $VE = 1 - IRR$

Simulation study

- Fitted a logistic regression model to identify significant covariates and calculated the vaccine efficacy with the below mentioned model with significance level 0.05 using R.

```
glm(formula = occds ~ trtgrp + age + comb, family  
    = binomial, data = simdata)
```

- occds- Occurrence of disease
- trtgrp - Treatment Group
- comb -Number of comorbidities
- Age – Age at the time of enrolment

Simulation Study -Inference

- Age, Comorbidities and Treatment Group are considered to be the significant factors affecting occurrence of disease.
- Checked the goodness of fit of the model with chi-square test and it showed the model is a good fit for the data.
- We got a high AIC value , indicating the goodness of fit of the model.
- Our model predicted that after vaccination the odds of getting diseased will reduce by 92.42%.



Thank You

