**Exercise leprosy in Comoros:**

Leprosy is a chronic infectious disease with a long incubation period. With the exception of a few animal species such as 9-banded armadillos, leprosy is assumed to be an anthroponosis, a disease with a human reservoir. The island of Anjouan in Comoros is highly endemic for the disease and we intended to explore whether providing a single dose of Rifampicin to close contacts of leprosy patients could help in curbing the spread of the disease. The dataset we will use in this exercise is shown in figure 1 below, it can be found in ‘leprosy\_comoros.csv’.

Text

Description automatically generated with medium confidence

Figure 1: Leprosy Comoros

The data shown were collected during two door-to-door surveys in which entire villages were screened for leprosy at baseline, in 2019, and again one year later. The data are in wide format and there are four variables, ‘PEP\_19’, ‘dist\_cat’, ‘pop’ and ‘inc\_case’. Contacts of leprosy patients detected at baseline were provided with a single dose of rifampicin as post-exposure prophylaxis (PEP). In some villages PEP was distributed to anyone living within 100 meters of an index case, in other villages only household members received PEP. There were also villages in which no PEP was provided. If ‘PEP\_19’ equals 1, it means the person received PEP in 2019. Variable ‘dist\_cat’ shows the distance to the nearest other person with leprosy at baseline (index case), either this is a household member, a neighbor within 25 meters, or a neighborhood contact at 25-50 meters, 50-75 meters, 75-100 meters, or beyond 100 meters. Variable ‘pop’ shows the number of persons in any category, variable ‘inc\_case’ shows the number of new cases that arose between the baseline survey and the survey in 2020. The index cases themselves (cases at baseline) were excluded.

Question 1:

Use the ‘aggregate’ function in R to determine the total population and how many cases arose over the follow-up period of 1 year? (First create a constant with a value of 1)   
What was the incidence rate?

Question 2:

What can you say about the effect of PEP, did it protect?

Question 3:

The data presented here mixes up 4 different study arms, in villages randomized to study arm 1, no PEP was provided, in study arm 2 PEP was provided to household contacts of index cases, in study arm 3 PEP was provided to entire villages and in study arm 4 PEP was provided to household contacts as well as others living in the same village and testing positive to a serological screening test but asymptomatic. As a result there may be an association between PEP and distance to nearest index case. If distance would also be associated with risk of becoming an incident case, distance could be a confounder. First check if distance is indeed associated with risk of becoming an incident case. Make a table with numbers of population, cases and risk ratios + confidence intervals for the 6 distance categories.

|  |  |  |  |
| --- | --- | --- | --- |
| **Distance** | **Population** | **Cases** | **RR(95% CI)** |
| Same HH |  |  |  |
| Neighbor < 25m |  |  |  |
| 25-50m |  |  |  |
| 50-75m |  |  |  |
| 75-100m |  |  |  |
| 100+m |  |  |  |

Question 4: Now that you have shown that distance is clearly associated with risk of disease, fit a poisson model that includes distance and PEP. What happens to the effect of PEP?

Question 5: Assume we are fitting a significance based classic model. Is the model with both PEP and distance better than the model with only distance?

Question 6: Check the dispersion parameter, is it far removed from 1? Is there a problem with the condition that mean should be more or less equal to variance?

Optional: Even though it is not necessary in this case, try to fit a negative binomial model. Does the result differ?