**GEE model, outcome and summary**

**My understanding on what statistical reviewer suggest:**

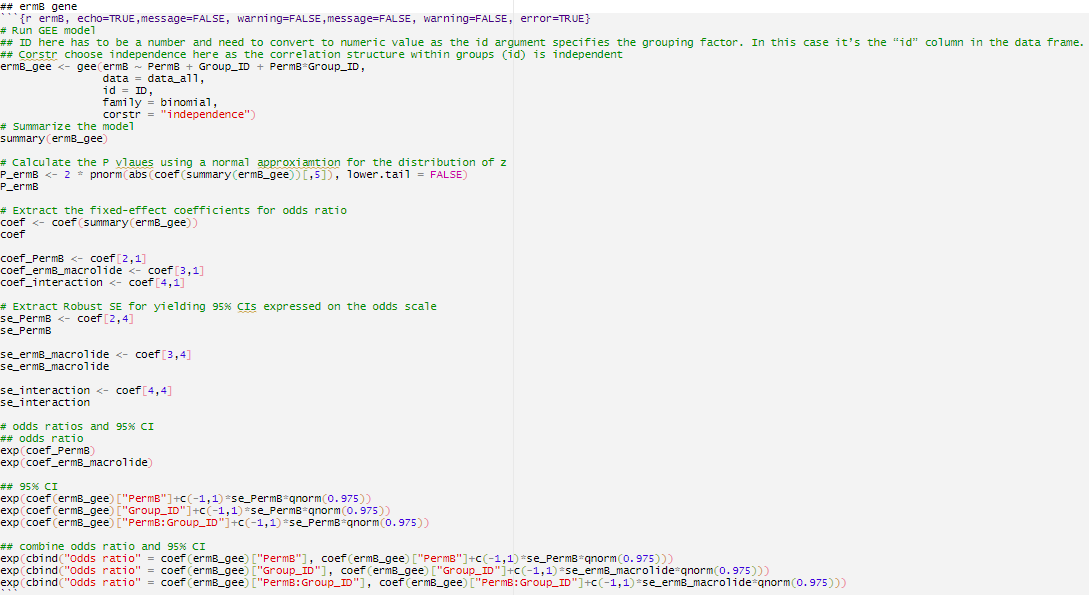
1. One GEE model: address both co-carriage question and transmission risk question
2. He emphasized one model and more efficient, which means run one time, otherwise it will not be more efficient than our models

**Coding system**

1. ​DV: detection of each gene in close contact (1/0)
2. IVs:
   1. \*First IV: gene detection in patients (1/0)
   2. Second IV: macrolide exposure (1/0)
   3. \*Third IV : macrolide exposure \* gene detection in patients (1/0)

\* indicates the IV that could answer our two questions

**Scripts example (Take ermB as an example)**



**Outcome**

1. Patient effect on detection of resistance gene in close contact
2. Macrolide effect on detection of resistance gene in close contact
3. Patient and macrolide effects on detection of resistance gene in close contact

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Resistance gene | Patient effect | | Macrolide effect | | Patient and macrolide effects | |
| Odds ratio | P value | Odds ratio | P value | Odds ratio | P value |
| *erm*(A) | NA | NA | NA | NA | NA | NA |
| *erm*(B) | NA | NA | NA | NA | NA | NA |
| *erm*(C) | 8.5  (0.4-163.9) | 0.16 | 1.7  (0.1-19.1) | 0.68 | 0.6  (0.1-6.9) | 0.78 |
| *erm*(F) | 1.7  (0.4-7.6) | 0.50 | 0.1  (0.01-0.8) | 0.026 | 7.0  (1.0-50.6) | 0.085 |
| *mef* | 1.3  (0.3-6.9) | 0.75 | 0.3  (0.05-2.0) | 0.21 | 5.5  (0.8-36.7) | 0.12 |
| *msr*(A) | 1.8  (0.4-8.2) | 0.43 | 1.3  (0.4-3.9) | 0.62 | 0.8  (0.3-2.4) | 0.85 |
| *msr*(E) | 1.1  (0.3-4.5) | 0.87 | 0.8  (0.2-3.5) | 0.73 | 0.7  (0.2-3.3) | 0.74 |
| *tet*(M) | NA | NA | NA | NA | NA | NA |
| *tet*(O) | 1.7  (0.4-7.6) | 0.50 | 0.7  (0.1-3.5) | 0.69 | 1.6  (0.3-7.8) | 0.64 |
| *tet*(W) | NA | NA | NA | NA | NA | NA |

NA: Cgee: error: logistic model for probability has fitted value very close to 1.”

This means there is not enough variance between groups to compute the outcome (e.g. majority of them are 1 or 0)

**Explanation of outcomes**

1. Patient effect on detection of resistance gene in close contact
   1. This is similar to our co-carriage question
   2. However, we cannot separate them into macrolide group or non-macrolide group, this is because macrolide exposure is also an IV in this model, if we separate them, we have to run multiple GEE models, then the P values and odds ratios are exactly same with our original model
2. Macrolide effect on detection of resistance gene in close contact
   1. This is question we did not include in our original manuscript
   2. Could be counted as an advantage
3. Patient and macrolide effects on detection of resistance gene in close contact
   1. This is similar to our transmission risk question
   2. If there is a significant result, it means the case where patient carried the gene (1) x macrolide exposure (1) is more likely to affect close contacts carried the gene, than other cases. Other cases here include:

A. 0 x 1: patient do not carried the gene (0) x macrolide exposure (1)

B. 0 x 0: patient do not carried the gene (0) x no macrolide exposure (0)

C. 1 x 0: patients carried the gene (1) x no macrolide exposure (0)​

* 1. The outcome can be translated to: "while the patient carried the gene, the macrolide exposure will significant affect close contacts carried the gene”. In other word, it tells us macrolide exposure will significant affect transmission case (1-1). Also, the significance could also be translated to " while there is a macrolide exposure, patient carried the gene will affect close contacts carried the gene". How to translate the results depend on your research question.

1. PROs and CONs of this GEE model

PROs:

* 1. It is efficient because we previously used two models
  2. It also answered another question we did not include in our manuscript, which is “Macrolide effect on detection of resistance gene in close contact”

CONs:

1. It did not answer exact same questions, our first question aims to know the co-carriage in different treatment group, however, it only give us the overall effect of detection in patient on detection in close contact
2. 0-0 case was included in the model, which will affect the outcome of macrolide effect on transmission risk; We could remove 0-0 case prior to GEE, but then it won’t be able to answer our first co-carriage question. We want to include 0-0 in our co-carriage model this is because we only want to know whether detection in close contact is dependent/independent of detection in patients.
3. Despite efficient a little bit, the GEE model, especially the interaction analysis, is hard to understand and is hard to be translated to clinical messages, not even to say the clinicians will understand the message we delivered
4. More strict for variances between groups, if there are less than 3 variances between groups, then it cannot compute the results, will tell you: “Cgee: error: logistic model for probability has fitted value very close to 1.”