**Ny Ony**

Dear Ny Ony,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. If you do not hear back from one other instructor before the end of the month, please let me know, as I will need to remind them to send along the feedback.

Overall

You have a fascinating research topic, and I really enjoyed learning from you in the class. Your slides are visually appealing and clear. However, they do not acknowledge other students who should have reviewed them and also listened to your presentation, as was a requirement in the assignment directions. It is also a bit unclear from your first slide why you have these two disparate questions—it would be helpful to define which is mechanistic and which is statistical and maybe provide a bit of background on the broad theme of your research.

Statistical

Your statistical model needs some work. You say that the distribution is binomial, but your response variable is the number of visitor species. A binomial response should be 0 or 1—for a number of species, this is most likely a Gaussian distribution or could be gamma or lognormal or negative binomial depending on the shape of the data. I’m also a bit confused by you independent (X) variables. I think you meant to write that your X variable was forest quality, meaning either disturbed or undisturbed—you should not have two predictors that are x1=disturbed and x2=undisturbed. Also, it looks like you have 10 plots with a bunch of other information about them, and I wonder what kind of structure you used to set them up? If, for instance you have a matched design of 10 disturbed sites matched with 10 udisturbed sites, you likely will need some nested random effects, and you might want to include some other environmental descriptors of the plots in your predictor variables.

Mechanistic

I love your mechanistic model! It is extremely simple and easy to understand. I have no edits to make to this, other than to challenge you to think about how you might write the equations for this system. Great work!

Next steps

These look like good next steps. The topics are pretty different, so they feel like they may end up being two different papers to me. You can actually build this mechanistic model and start to explore these relationships without actually having any data to start. For your step about combining data you already have, how do you intend to define what constitutes a disturbed vs. undisturbed site? By including some environmental variables in your predictions, you might be able to avoid having to make some arbitrary decisions about how to classify these landscapes.

Great job!

Cara

**Miroso**

Dear Miroso,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. If you do not hear back from one other instructor before the end of the month, please let me know, as I will need to remind them to send along the feedback.

Overall

Great job on these slides. They have a common background and are nicely organized and easy to follow. The overall title is a little awkward as written in English—I might suggest you change it to “COVID-19 case outcomes under differing therapeutic regimens” or something like that. I’m glad you were able to share your slides and receive feedback from so many different instructors. Which of these acknowledgements reviewed your text vs. listened to your presentation?

Statistical

This statistical model looks pretty good. For the family, the term is actually a ‘binomial’ distribution when there is a binary (0 or 1) outcome. A bit confusing but that is the proper terminology. You might also want to think about adding a star between your two predictors to also assess interaction effects. If written as “therapeutic\_outcome~age\*gender”, R will test the independent effects of age and gender on therapeutic outcome, as well as the interaction of age with gender on outcome. Finally, you don’t currently have any random effects listed (e.g. different sites or repeated infections in the same individuals), so you don’t actually need to use the glmer() function. The glm() function alone will suffice.

Mechanistic

This model looks pretty good as well. You will want to add the dashed arrows to show the impact of the infected population on the transmission rate for COVID-19 (and those in both infected classes will be able to influence the transmission rate for the opposite class). You could also think about reordering this model to pass from S to I and then either go directly to R (for those not on corticotherapy or into IC (for those in corticotherapy) before reaching R. This would allow you to explore the impact of changes in the rate by which patients are introduced to therapy on their outcomes.

Next Steps

These are fantastic next steps and all very achievable. Once you have your data in hand, please keep in touch if you have additional questions about how to actually implement these models. Great work!

Best,  
Cara

**Césaire**

Dear Césaire,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. I see that you’ve already had feedback from Andres, so this is the last you will hear from us for awhile.

Overall

I agree with Andres that the common background for the slides makes it easy to follow your progress. Nice work! For your background, it might be nice to include a sentence or two that highlights that plague is a big problem in Madagascar (which is unusual for the rest of the world) in case you have a broader audience.

Statistical

I agree again with Andres on simplifying the question. Also, as currently written, you don’t need to separately include predictor variables of rodent abundance and season and also rodent abundance\*season. The “\*” symbol in R will simultaneously test each individual predictor’s independent association with the response variable, as well as the interactive association. If each site was sampled across multiple seasons, you could also instead think about moving the “season” predictor to be nested within site as a sub-random effect. It’s great that you already have so much data to address this question!

Mechanistic

This model looks pretty good. A few suggestions: as Andres mentioned, I don’t think it is necessary to include the “survival rate” since survival is simply 1-death rate and you could always calculate one from the other. Instead, I might suggest that you have a background mortality rate that is the same across both the Susceptible and Infected boxes and then an additional infection-induced mortality rate for the Infected boxes. You also have a dashed arrow that shows rodents passing from Infected back to Susceptible, and I am not sure what you are trying to demonstrate there. The dashed arrow signifies an influence, so it should relate a population in one box with a rate somewhere else in the model, rather than another box. If you instead want this to represent a rate itself, meaning a flow of rodents from the Infected class back to the Susceptible class (meaning they survive infection and can then be infected again at a later timepoint), this should be a solid arrow.

Next steps

These look logical and feasible. I appreciate that you recognize that this is work-on-progress and that further practice and review will be needed. Very wise! It is also great that you plan to read the literature more—a lot can be learned from reading what others have done in a system.

Great job!

Best,

Cara

**Nambininanahary**

Dear Nambinina,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. If you do not hear back from one other instructor before the end of the month, please let me know, as I will need to remind them to send along the feedback.

Overall

Great job on these slides, and thanks so much for coming to the office hours to seek help even when you could not make the symposium. These are nicely organized and you do a great job explaining the relevance of the research topic and defining your two different questions clearly. The acknowledgements were supposed to be targeted towards sharing your homework with other students in the course, but it is great that you are getting feedback from all of these difference groups regardless.

Statistical

This statistical model looks quite good! However, you can simplify the question a bit to beter match the data you have on hand and the way you have structured your model. You don’t actually measure prevalence (the proportion infected) which is the response variable you highlight in your question. Instead you measure occurrence, so your question should be changed to reflect that. I would suggest writing it as, “What is the relationship between the occurrence of CSF in domestic pigs and the presence of infected bush pigs in the surrounding region?” Beyond that, your model looks great!

Mechanistic

This also looks great, and the question is perfect for a mechanistic modeling approach. My one question for you would be that you show a dashed line from the latent domestic pigs (LD) to the transmission rate for the bush pigs (betaB): do you really think that domestic pigs are responsible for a sizeable amount of transmission back to the bush pigs? If not (if it usually transmits only bush to domestic pigs and rarely in the opposite direction), then you might want to start with the simpler model that leaves this transmission route out. If, by contrast, you think this is really important, then you can leave it in, and I would also ask if you should then add a dashed line from ID to betaB as well? Only if you think it represents a sizeable and important-to-include route of transmission. Additionally, it’s great that you started to write out the equations, though they were not required for this assignment. Since you started, however, how would you write the equations for the other state variables as well?

Next steps

These look great! This is an impressive project and a really thorough research theme. Will you be assessing infection by serology only or also by PCR? Will the data be collected for both domestic pigs and bush pigs or only domestic pigs? And, finally, how do you plan to distinguish individuals in the L and I classes in your model? Nice work!

Best,

Cara

**Zina**

Dear Zina,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. If you do not hear back from one other instructor before the end of the month, please let me know, as I will need to remind them to send along the feedback.

Overall

These slides look great—the common background is very professional, and you did an excellent job getting feedback from so many other students in the course. The opening slide was supposed to include the statistical and mechanistic questions to provide a bit of a road map for where the presentation was heading next. However, you gave a very clear presentation, and no one had any trouble following along.

Statistical

This model design looks excellent. I really like that you included the nested random effects of household within fokontany. As it is currently written, I do not think you will be able to differentiate between hypothesis 2 and 3, however, since your predictor is only the presence or absence of non-communicable disease and therefore won’t look any different for a person who has high blood pressure vs. a person who has diabetes. You could improve this in a couple different ways. If each person can only have one type of non-commmunicable disease, thenou could make X2 a categorical predictor that simply lists the type of disease (e.g. diabetes or high blood pressure). If, as I expect, it is possible to have both diabetes and high blood pressure, then you can just add another predictor variable—so you would have X1, X2, and X3 for STH, diabetes, and high blood pressure. If you want to explore the interaction of STH infection with these noncommunicable diseases, then you could write the R Code as “Y~X1 + X2+X3 + X1:X2 + X1:X3”. The “:” symbol tests the association between the interaction of STH and diabetes for example on infant birth weight.

Mechanistic

This also looks very good, especially assuming you have data on both populations with and without MDA where you can fit this model and estimate different parameters. You could also experiment with a different type of question that asks instead “How does MDA influence the prevalence of STH through time?” and you could then have the introduction of MDA at various timepoints in a time series and see the impact on the prevalence. One thing you likely want to add is an arrow from R back to S in both of your models because I think it is possible to get reinfected again after recovery. If it is possible for this to happen very fast, then you may not even need an R class—you can have the recovery rate take individuals directly from chronically Infected back to susceptible. Depending on the helminth that you are modeling and its route of transmission, it can sometimes also be appropriate to model an environmental reservoir for the helminth as well.

Next Steps

Good plans here! For the mechanistic model, it can sometimes be challenging to collect data on the rates of process but easier to collect data on the state variables. You can then actually estimate the rates by fitting the model to the state variable data.

Great work!

Cara

**Vero**

Dear Vero,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. I see that you’ve already had feedback from Andres, so this is the last you will hear from us for awhile.

Overall

I agree with Andres that your slides look great. You did a very nice job providing background information on the opening slide, as well as preparing us for the focal questions ahead. Nice work soliciting feedback from other E2M2 students as well.

Statistical

This looks pretty good, but I do think you could improve it some. As Andres mentioned, it would be helpful to have a bit more information about the sampling design to better prepare the model: for example, if you have the same sites repeatedly sampled over many years, you might need some kind of random effects structure, which would require you to use glmer(). My main question relates to your response variable: if you are testing (0,1) whether or not the sequevar is identified in a given site, then I think you will also need a predictor (x) variable that is the type of sequevar. Then, because you are simultaneously testing many hypotheses (e.g. the presence/absence of several different types of serovars), you will also need to do a correction (often called a Bonferroni correction) to defend against type I error. I’m happy to answer any additional questions you might have on that front downstream.

Mechanistic

This also looks great and is certainly a great framework for mechanistic modeling. My main question would be about the influence arrow you drew from the Tolerant plant box to the birth rate of the susceptible plants—are you suggesting that tolerant plants either increase or decrease the birth rate of the plant population as a whole? I don’t quite understand this bit. Additionally, the arrow for the infection-induced death rate (alpha) is facing the wrong direction: it should point outward from the infected class box. Like the natural death rate (mu), infection-induced mortality provides a mechanism by which Infected plants leave the population.

Next Steps

These next steps look appropriate! Your work is very impressive and really quite feasible for publication. Keep in touch if you have additional questions in the future!

Best,  
Cara

**Pierro**

Dear Pierro,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. If you do not hear back from one other instructor before the end of the month, please let me know, as I will need to remind them to send along the feedback.

Overall

Nice job overall on your slides…You were technically supposed to have 4 slides, with a ‘Next steps’ slide as the last one, so that is missing. You are also missing the acknowledgements to other students who should have reviewed and offered feedback on your slides and presentation. It would also be helpful to have a bit more background here about some of the threats to corals in Madagascar and the motivation for this work.

Statistical

In general, this looks really good. It seems that your question at the top is cut off, however. You wrote “What is the factor influence juveniles abundance” …Maybe this was supposed to be “What factors influence juvenile coral abundance in southwestern Madagascar?” Other than that, though, your model structure looks excellent. Typically, the random effects are written as the last item mentioned in the predictor variables, but this is just stylistic. It is fine to put them in the middle as well. Additionally, there is a chance that a Gaussian family won’t work for your data if you plot them and discover that they are not normally distributed. In those cases, you can think about log-normal or Gaussian or negative binomial distributions as useful alternatives.

Mechanistic

This model looks excellent! There is a lot of literature out there on Lefkovitch or Leslie matrix population models, even a whole book by Hal Caswell. This system lends itself to this modeling structure really nicely, and you’ve done a great job representing it. My only criticism would be to say that the arrow from Adult to Recruit should probably instead be a dashed arrow (an ‘influence’ arrow) that points to a recruitment rate, since it is not technically true that the adults become juveniles and since the adult population does not lose numbers as the recruits grow. Your equations look excellent, though—nice work.

Next Steps

These are missing but the work looks right on track!

Best,

Cara

**Parfait**

Dear Parfait,

Great work at E2M2 this year. It was a pleasure to have you as a student! Please see below for my final feedback, and don’t hesitate to let me know if you have any outstanding questions. If you do not hear back from one other instructor before the end of the month, please let me know, as I will need to remind them to send along the feedback.

Overall

Nice work on these slides! You did a great job of providing background and motivating the question about why we care about this research subject. I also really like the scale bar at the top that helps the viewer keep track of where we are in the presentation at any given time. It might be good to add identifiers next to your two questions on the opening slide to indicate which one is statistical and which one is mechanistic. Also, you were supposed to include acknowledgements to a few other E2M2 students who should have reviewed your slides and also listened to a practice of your presentation—I would encourage you to get feedback from peers, since so many are working on plague in Madagascar.

Statistical

Nice job on this! Everything looks to be in proper order. You actually have an option in your nesting in the random effects structure: if you are interested more in differences between urban and rural, you could have that first and then nest each site within the two categories. Conversely, if you want to highlight sites between each other, you can leave the structure as it is currently worded.

Mechanistic

This is an excellent model. I love that your really distinguished the routes of transmission with the influence arrows: fleas can only get infected by biting an infected rat, and those infected fleas go on to transmit to other rats and to humans—but humans also can get infected directly by each other, via pneumonic transmission. My question for you would be: do you have any data, or do you plan to do experiments to try to recover a value for the parameter of birth rate dependent on temperature?

Next Steps

These look great! I really like that you plan to do a literature review first—so much has been done on plague in Madagascar that it is important to find out where the gaps are. Do keep in touch if you have questions in the future. Great work!

Best,

Cara