1. Applicability and Generalization

The deep learning model performs well on simulated data, but how well does it generalize in real-world applications? Is it applicable to various pathogens or epidemiological models, and how would it perform with novel or emerging pathogens?

Discussion: Explore whether the model requires customization for different diseases (e.g., with varying transmission routes, incubation periods) or if the neural network can be retrained for different settings. Also, discuss its potential applicability to other fields like population genetics or macroevolution.

2. Comparison with Traditional Methods

Question: What are the specific advantages and disadvantages of this deep learning approach compared to traditional methods like Maximum-Likelihood Estimation (MLE) and Bayesian inference (e.g., BEAST2)? Are there situations where traditional methods might still outperform deep learning?

Discussion: While the deep learning method shows significant improvements for large datasets, are there cases where traditional methods, especially for smaller datasets or simpler models, might provide more reliable or stable results? Furthermore, how do traditional methods fare in handling uncertainty, where they provide posterior distributions instead of just point estimates?

3. Model Interpretability

Question: Does the "black box" nature of deep learning models pose problems in epidemiological inference? How can we interpret the predictions from the deep learning model in practical applications?

Discussion: Consider how to improve the interpretability of complex deep learning models, perhaps by combining traditional statistical methods or using feature importance analysis. Particularly, how can we ensure that the inferred epidemiological parameters are transparent and trustworthy when used in public health decision-making?

4. Dependence on Simulation Models

Question: Since the deep learning method relies heavily on simulated data for training, how would biases or inaccuracies in the simulation model affect the final inference? Could these errors be amplified in the results?

Discussion: Investigate the impact of simulation biases on model training and how they might affect real-world applications. Discuss ways to mitigate these effects, such as using more diverse or comprehensive simulations, and whether the model can still perform well with limited or imperfect simulation data.

5. Handling of Extremely Large Datasets

Question: While the deep learning method shows great scalability with large datasets, how will it perform as datasets grow even larger (e.g., trees with hundreds of thousands of tips)? Will the method maintain its speed and efficiency?

Discussion: Consider the scalability of the method for future ultra-large datasets. With the growing availability of pathogen genetic data (e.g., from COVID-19), will further optimizations or algorithm improvements be needed to handle the exponential growth in data size?

6. Computational Resource Requirements

Question: While the deep learning method offers fast inference once trained, the training phase requires significant computational resources, especially for large-scale simulated datasets. How can we balance the computational demands with model performance?

Discussion: Explore the potential limitations imposed by computational resources. Could the training process be optimized (e.g., through model pruning, low-precision computing) to reduce resource requirements? Would distributed computing or cloud-based infrastructure be necessary for real-world applications?

7. Uncertainty Handling

Question: Although the deep learning method is much faster for inference, is it sufficient in handling uncertainty? Can it provide posterior distributions for parameters like Bayesian methods, or does it mainly provide point estimates?

Discussion: Discuss how to combine Bayesian methods with deep learning to ensure efficient inference while also quantifying uncertainty. Could techniques like Bayesian deep learning or adding priors to the neural network parameters address this gap?

8. Performance on Real-World Data

Question: The deep learning model performs excellently on simulated datasets, but how well does it perform on real-world datasets? Does the difference between simulations and real-world data affect the reliability of the inference?

Discussion: Explore how the deep learning model trained on simulations can generalize to real-world pathogen transmission data. Would adjustments or cross-validation techniques be necessary to ensure that the model performs well on actual outbreak data?