Pregnancy Outcomes and Heat Stress

Alejandro Mantero

Naresh Kumar

Introduction

- Preterm birth is an issue affected 10.4% of births in Florida in 2020, it has been associated with a number of developmental issues throughout a child's life
- It was suspected that exposure to heat measured via heat index may play a role in tropical environments in exacerbating the risk of preterm birth during certain times of the year.
- In order to explore potential associations weather, census and health data were collated and then analyzed both via traditional regression methods (GEE) and more flexible machine learning methods (GBM via boosted regression trees)
- The aim is to determine where there is commonality and dissimilarity between the approaches which may indicate places that require further research, i.e., overlooked aspects due to misspecification of associations

Data

- This was a data exploration, initially of all the births between 2016 and 2020 though due to computational limitations and similarity in effects observed year-by-year, it was limited to 2020 for greatest relevance.
- This data was combined with the collocated weather data at the county level and heat index was calculated and lags of up to 45 days were calculated.
- This data was augmented with county level information from: https://github.com/JieYingWu/COVID-19_US_County-level_Summaries/, a repository of high quality preconstructed data from various US agencies

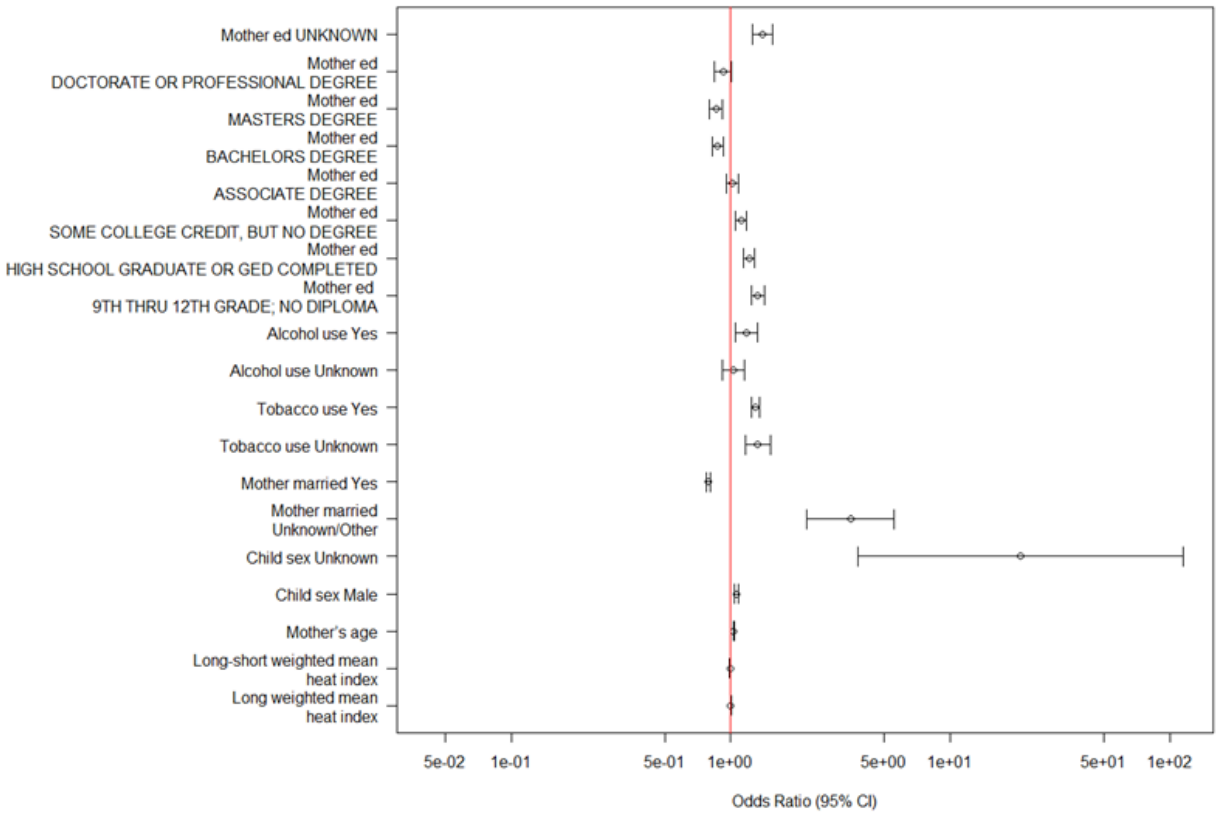
Methods

- Two approaches were implemented to represent possible different avenues and possible different discovered effects.
  - First, a GEE model was fit using mother level characteristics to determine associated factors, correlation by county were accounted for and an exchangeable correlation structure was implemented.
  - Second, a boosted tree model was implemented, optimization of the training parameters was performed on the training set. Important note, the goal of this model was not prediction but determination of possible nonlinear associations therefore no test set was involved, and optimization was on fit not MSPE.
- Due to class imbalance oversampling of preterm birth observations had to be utilized to facilitate model training.

What about heat shock?

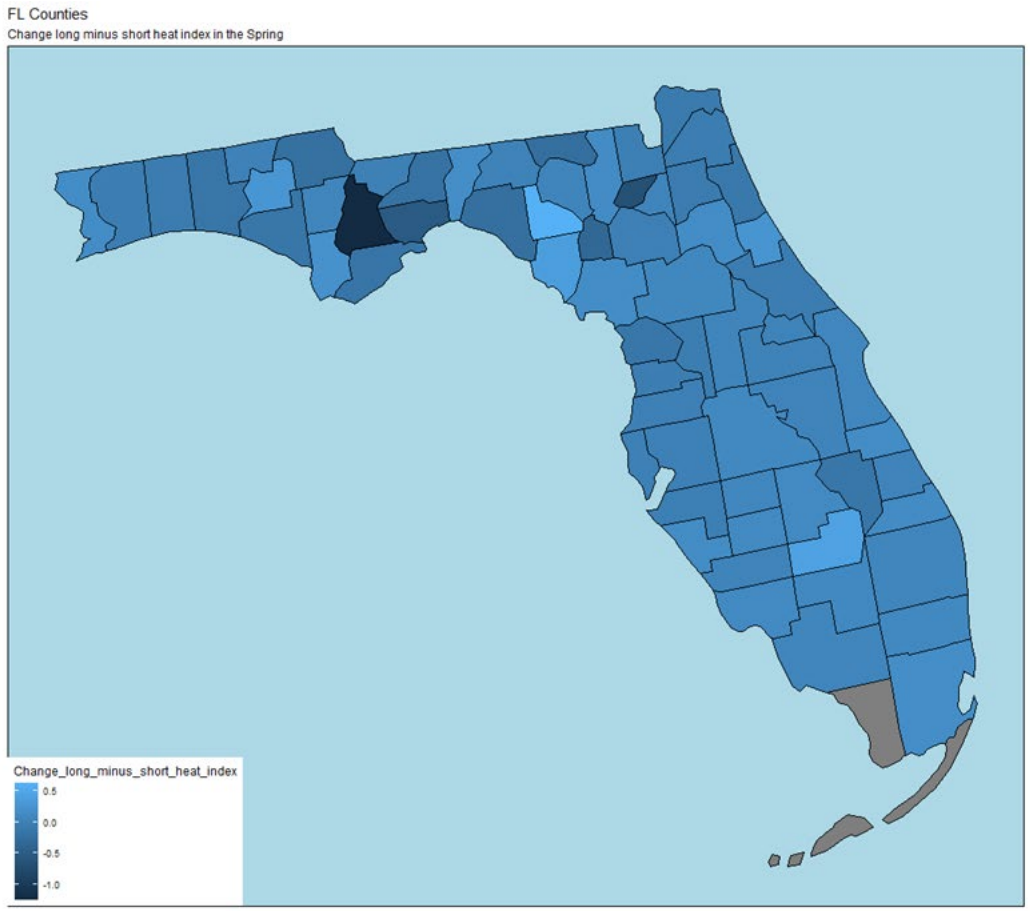
- Although we have measurements of heat index, we had two limitations that needed to be addressed
  - We could not include heat index for every lag or for even a small set because of multicollinearity
  - Sure, heat index might be associated but more likely is the change in heat index
- For this reason, we performed a data exploration correlating via Spearman, preterm birth with heat index and found that for lags 0-20 there was a positive correlation while 21-45, it was negative
- This gave us evidence of 2 eras and potential the near term was almost like shock while long term was adjustment
- This gave rise to a constructed variable of heat shock using the difference of the weight averages for the long and near term eras.

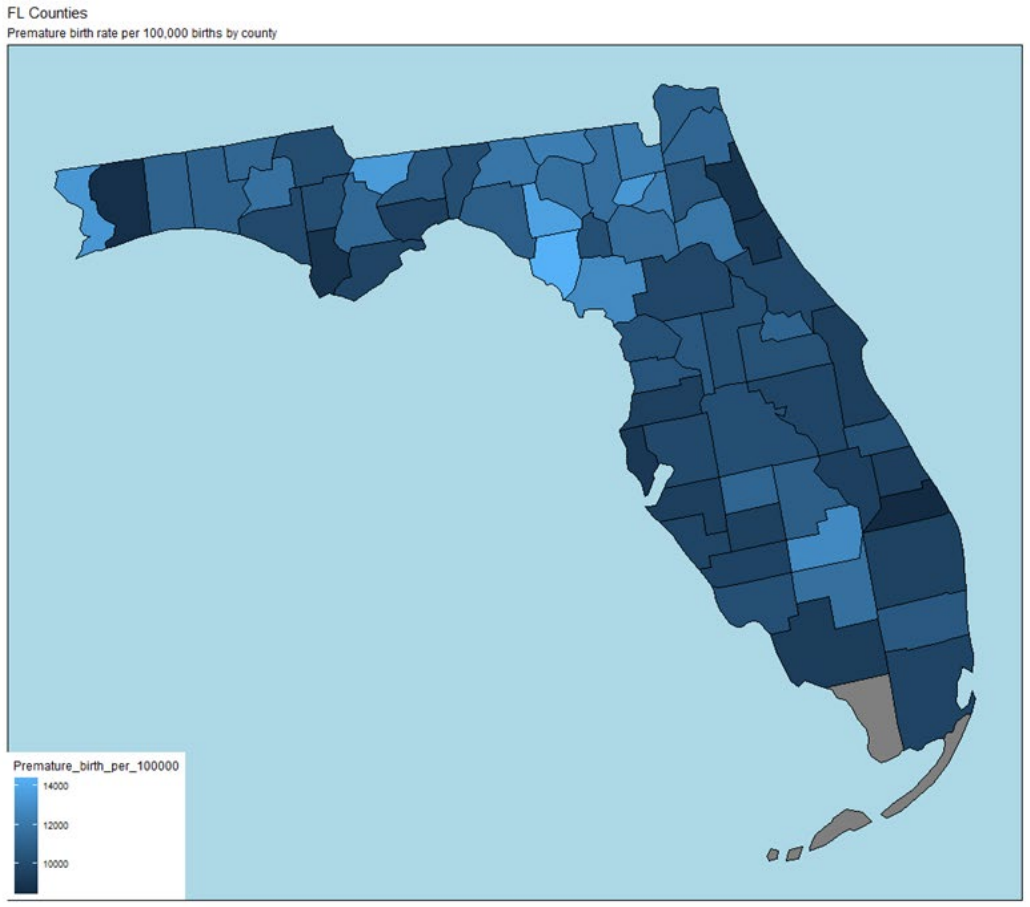
GEE model results



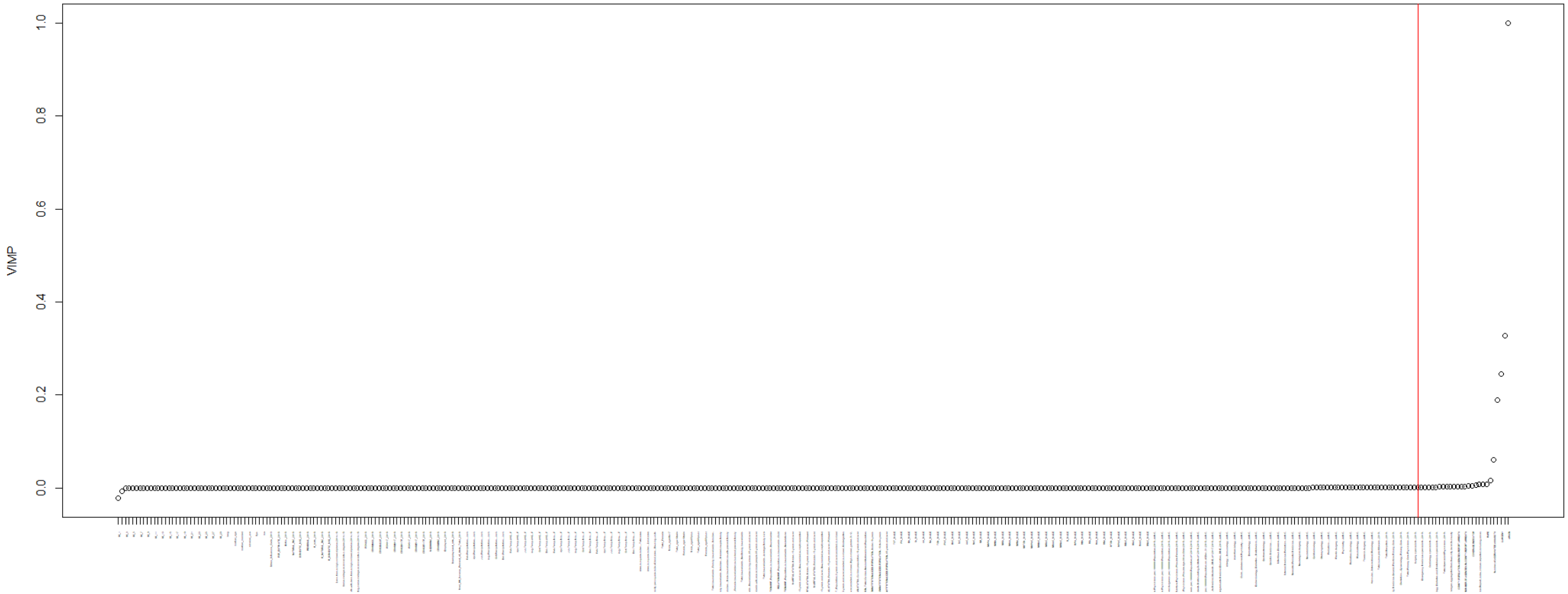
- Note the directionality of the heat shock variable
- It is negative so as short-term heat index is higher, we have increasing heat which corresponds with an increased risk of pre-term birth
- To put into context, if we take the average increase in heat shock between April and May for Miami this would translate into a 3% increase in risk.

Confirmation of heat shock

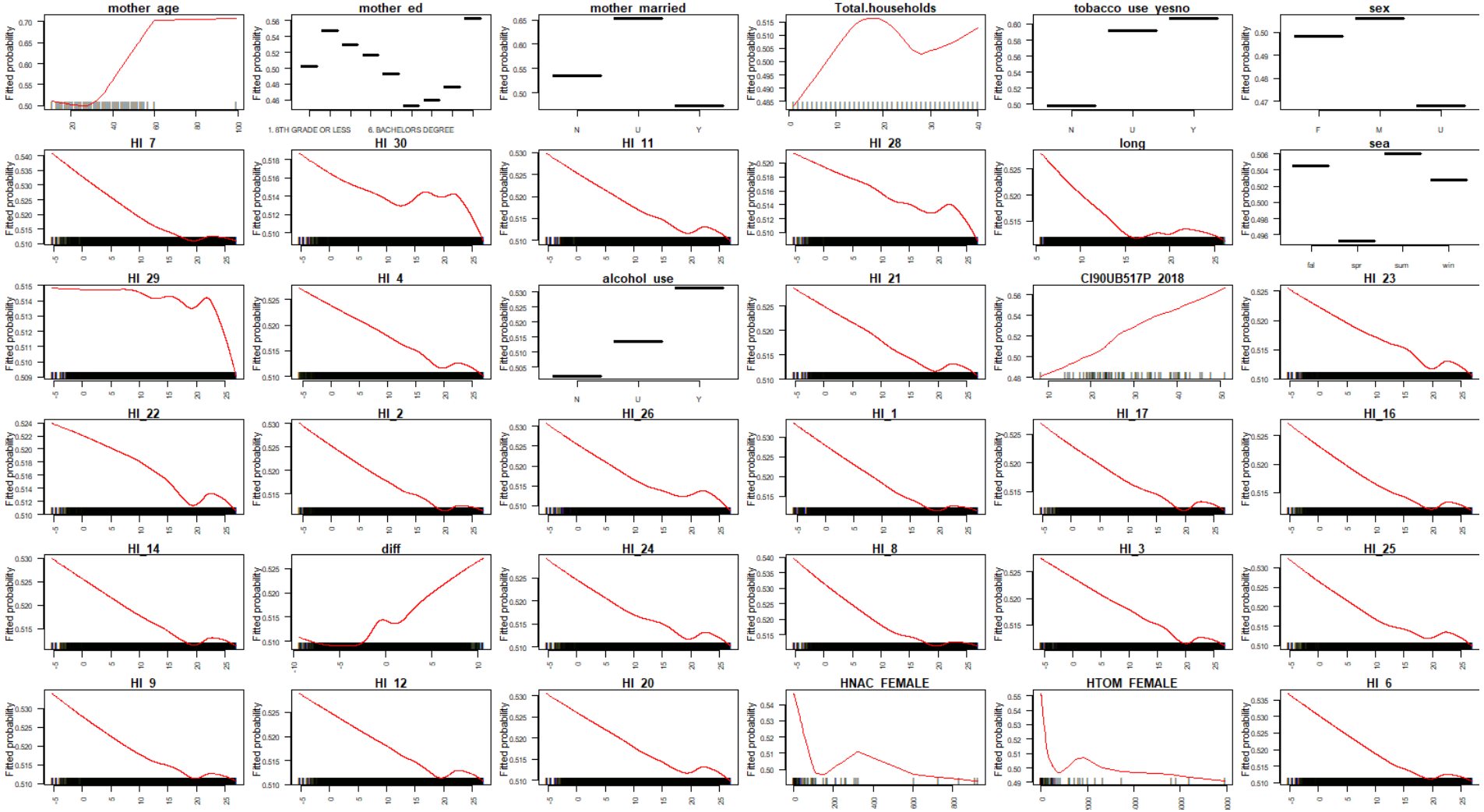




Boosted trees VIMP



Boosted regression tree marginal plots



Discussion

- It should be noted that the directionality of the heat shock variable was opposite between the two models
- This could be a result of having more variables at hand in the GBM model and some of the lag terms may explain enough of the heat shock variable to push it into the other direction
- Linearity may be a strong assumption that is forcing the directionality in conjunction with the variables available
- Also, of note is how both the GEE and GBM models characterize patients with missing data, they have a tendency of being more at risk suggesting a possible association between data collection reliability and outcome, potential indicator of health system quality.

Conclusion

- There appears to be some evidence in favor of an association of environmental factors with the risk of preterm birth
- Heat stressors may need to be mitigated especially in times where the prevailing climate changing, i.e., fall and spring
- Although the associations may look weak at present, when extrapolated to a population of millions, it is an appreciable number potentially affected
- Further research into the proper specification of heat stress in terms of timing and measurement seems to be warranted.

Thank you for your time

Any questions?