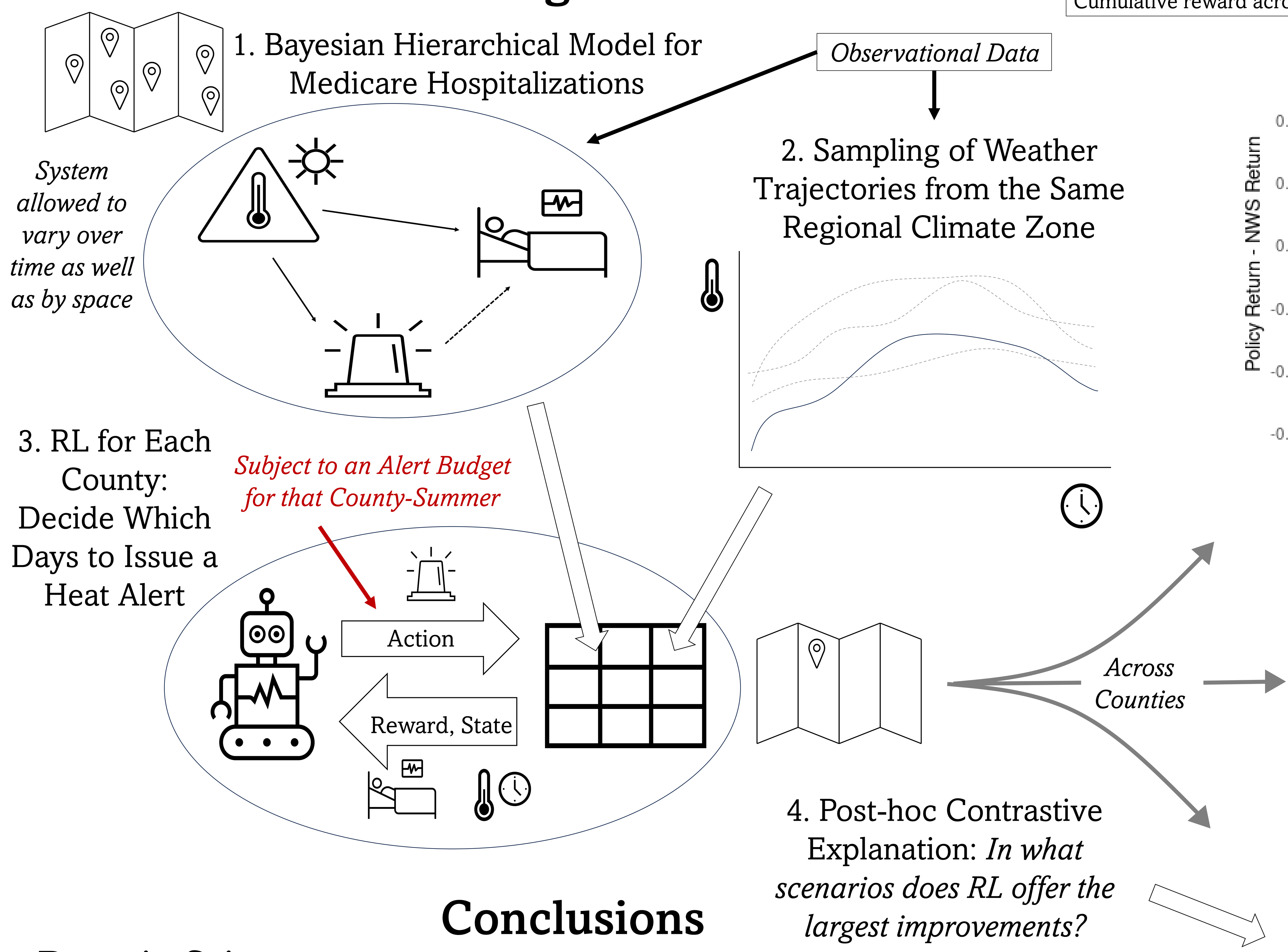
Optimizing Heat Alert Issuance for Public Health in the United States with Reinforcement Learning (RL)

Ellen M. Considine, Rachel C. Nethery, Gregory A. Wellenius, Francesca Dominici, Mauricio Tec

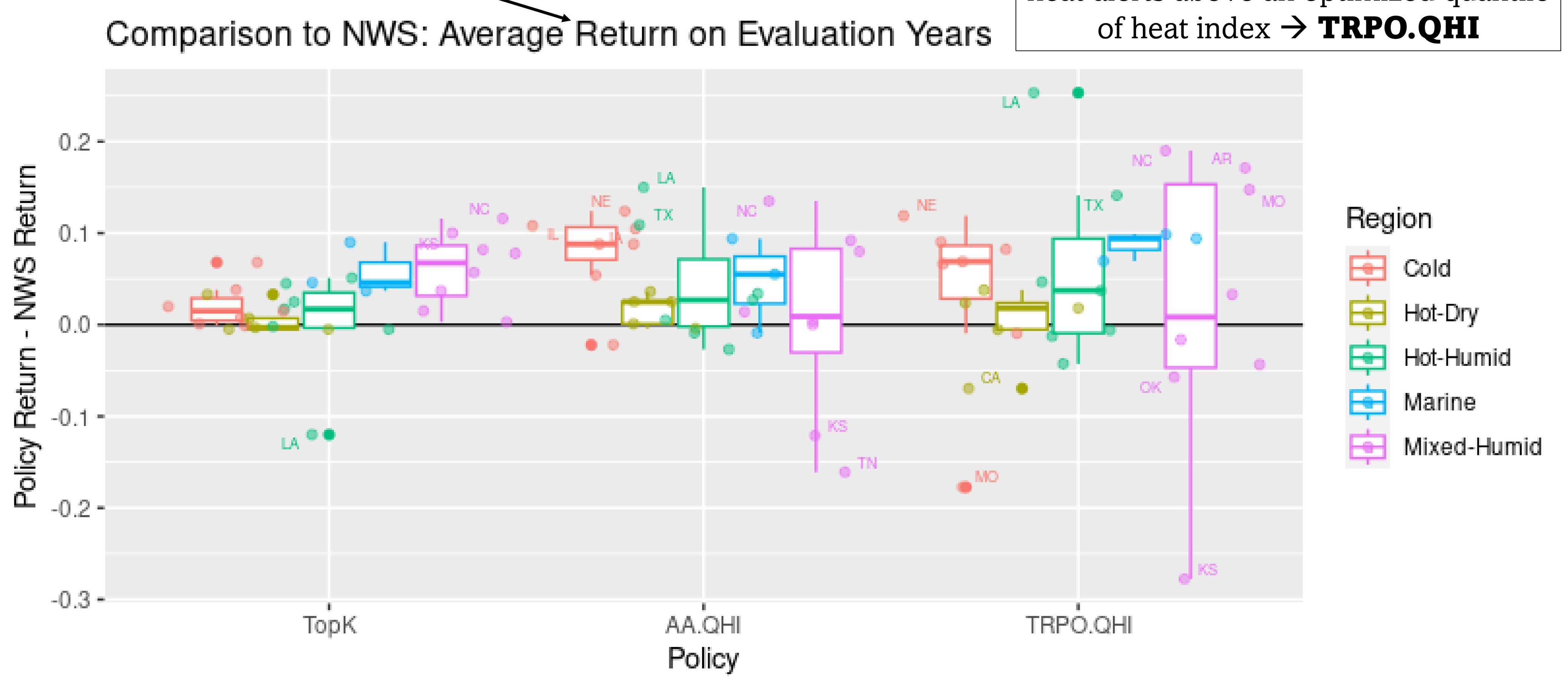
When should we issue heat alerts, accounting for geography, socioeconomics, and sequential dependence of alerts?

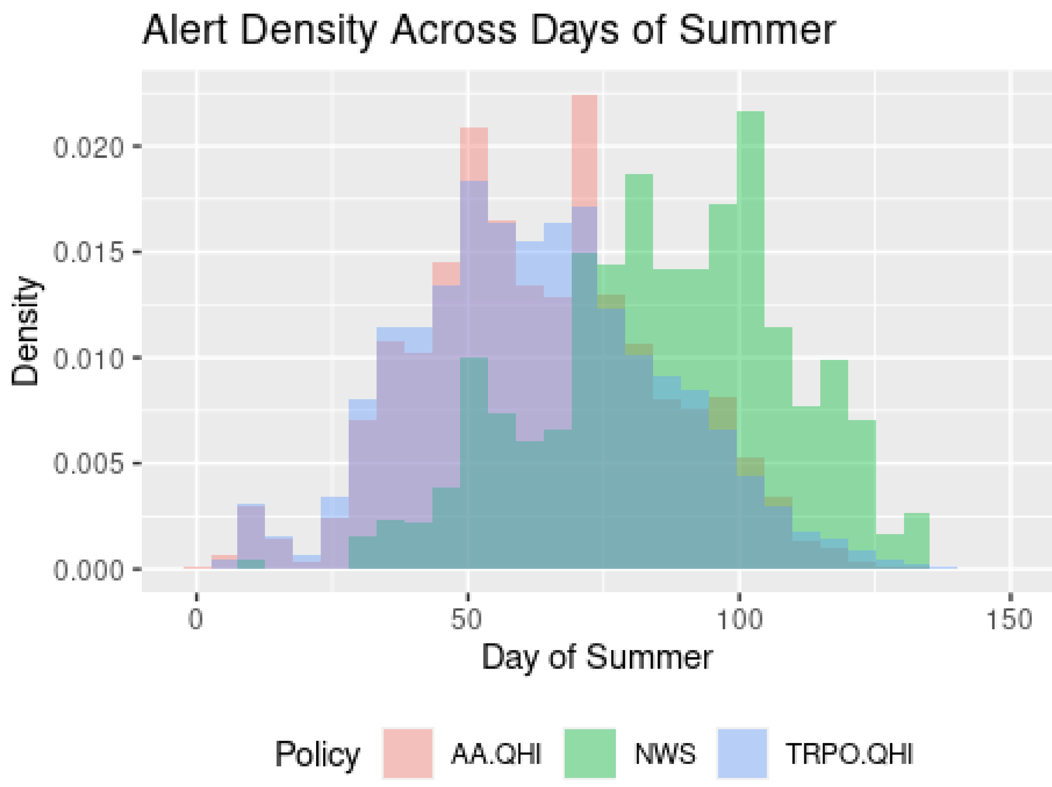
Methodological Framework

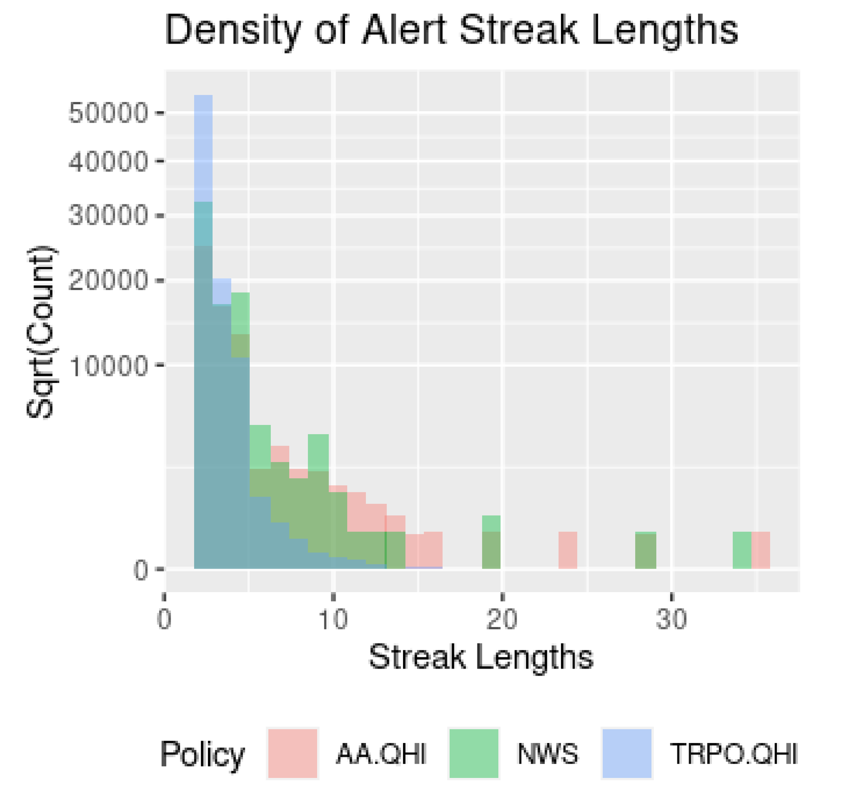


Results

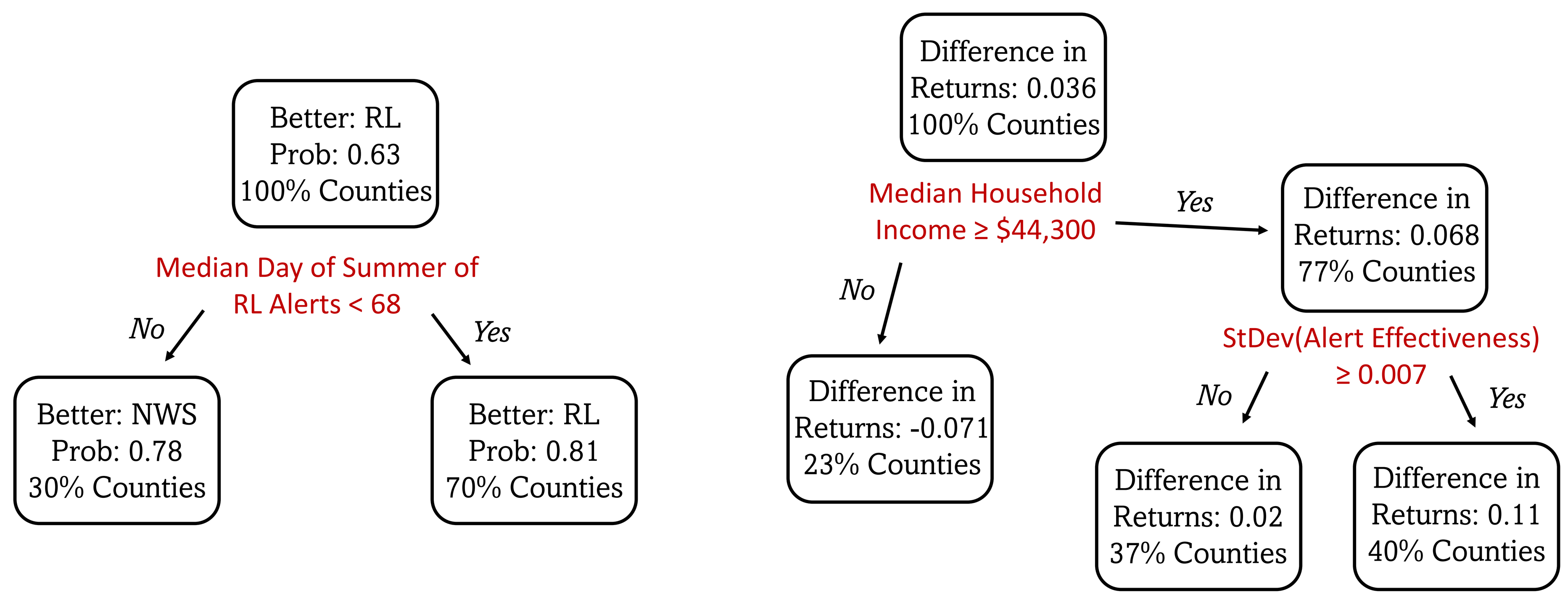
Our best RL model used the TRPO algorithm, constrained to only issue heat alerts above an optimized quantile of heat index → **TRPO.QHI**



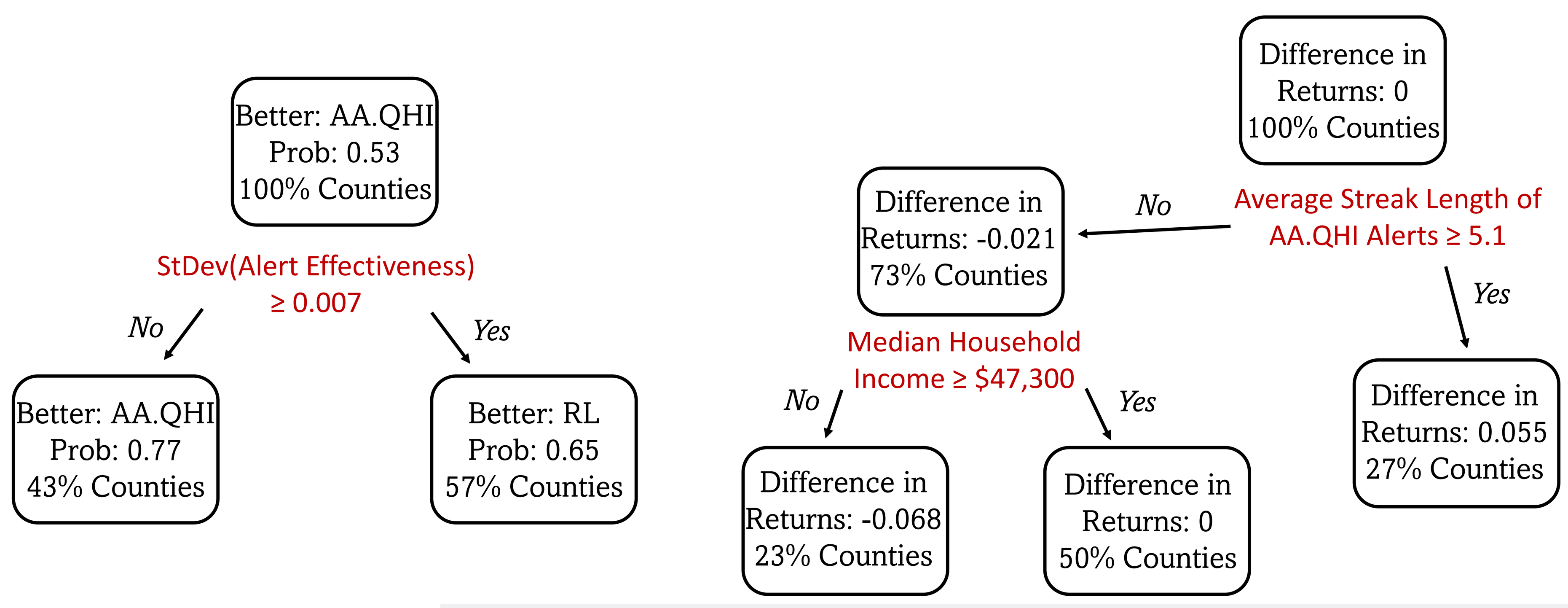




Classification (LHS) and Regression (RHS) Trees [CART] Comparing Best RL to **NWS**:



Classification (LHS) and Regression (RHS) Trees Comparing Best RL to **AA.QHI** (Always Alert Above an Optimized Quantile of Heat Index for that County):



Conclusions

Domain Science:

1. Must consider *not-obviously heat-related* hospitalizations to avoid mediation by more people seeking care after seeing alerts
2. Evidence of alert fatigue in our rewards (hospitalizations) model
3. Our counterfactual policies outperformed the NWS with statistical significance, but exhibited large heterogeneity across counties → "safe" policy learning could help
4. Intuitive insights about *where* RL offers greatest benefits: locations with more prolonged heat waves (e.g. high humidity), larger heat alert-health signal – especially earlier in the summer

Methods:

1. Off-the-shelf RL methods are inadequate to solve this problem
2. Our framework lays the foundation for sequential decision making in environmental health
3. Stochastic policy gradient RL performed better than value learning
4. Limitations: using fixed alert budgets, sensitivity of results to specification of rewards model, nontrivial uncertainty quantification

NSAPH Research Group

HARVARD T.H. CHAN SCHOOL OF PUBLIC HEALTH