National Center for Health Statistics

County-Level Estimates of Mortality and Natality Indicators from the National Vital Statistics System

Lauren M. Rossen, Ph.D., M.S.[1]
Diba Khan, Ph.D.[2]

[1]Division of Vital Statistics, National Center for Health Statistics
[2]Division of Research Methodology, National Center for Health Statistics

FCSM Research Conference

March 7, 2018

Acknowledgements

- **Co-authors and contributors:**
  Diba Khan
  Brady Hamilton
  Margy Warner
  Ashley Hirai
  Michael Kramer

County-Level Estimates of Natality and Mortality Indicators

❖ Natality

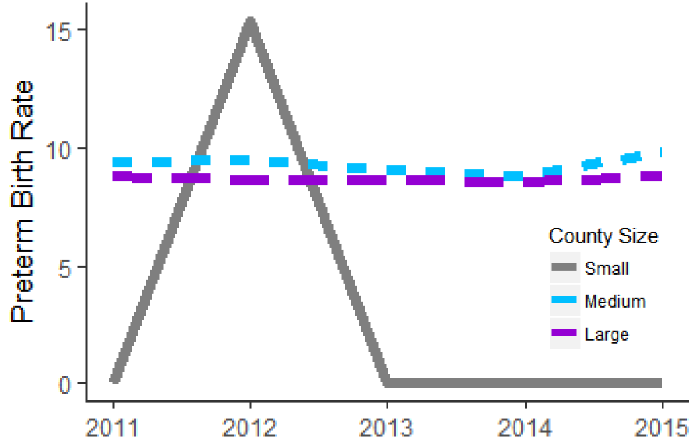
- Preterm birth
- Second and higher order teen birth rates

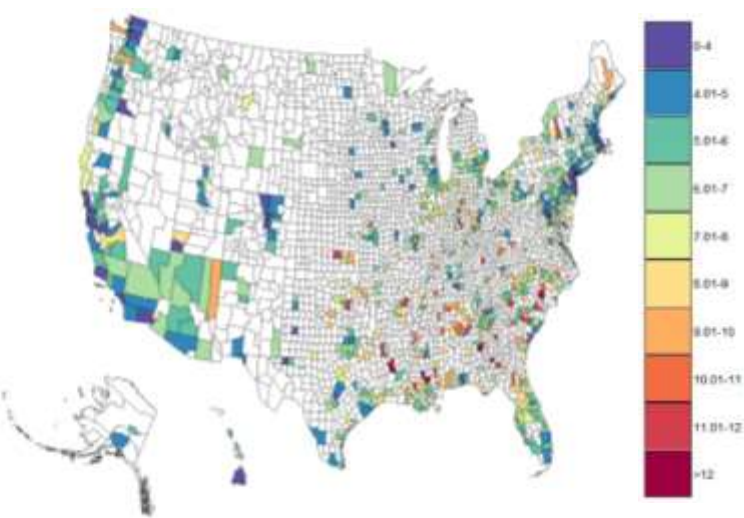
❖ Mortality

- Infant mortality

Rationale

Birth or death rates at the county level are often unstable ➔





🡐 Rates suppressed for counties with < 20 births/deaths

Outcomes from the National Vital Statistics System

Preterm birth (2013-2015)

- Percent of infants born before 37 completed weeks gestation
- Aggregated over 3 years

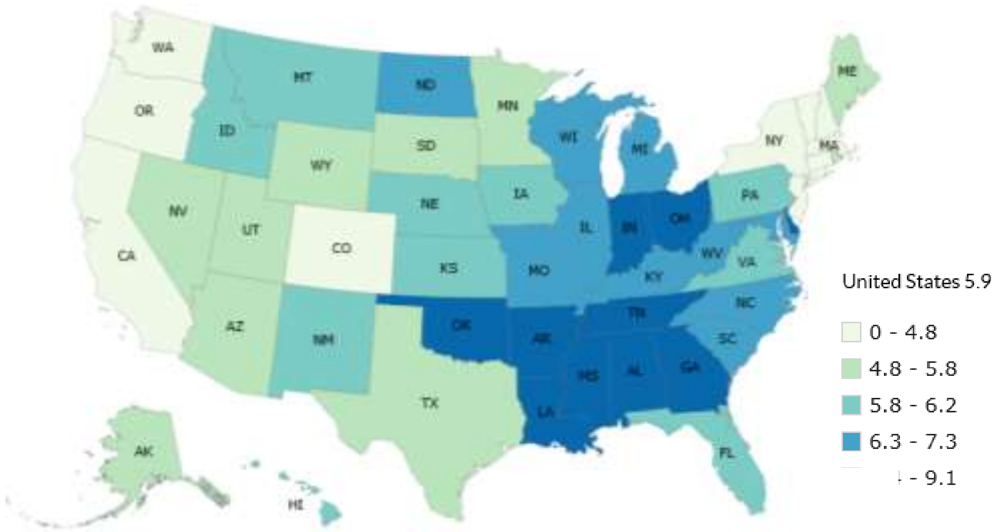
Second and higher order teen births (2007-2016)

- Repeat births to teen mothers
- Number of second or higher order births per 1,000 females 15-19 years
- Annual trends over 10 years

Outcomes from the National Vital Statistics System

Infant mortality (2013-2015)

- Infant (< 1 year of age) deaths per 1,000 live births
- Aggregated over 3 years



Methods

- Hierarchical Bayesian models
  - Integrated Nested Laplace Approximation (INLA) in R
    - Latent Gaussian models
      - Besag, York, Mollié (BYM) models
        - Spatial random effect, intrinsic conditionally autoregressive structure
        - Non-spatial random effect
    - Fast and flexible
      - Many ‘built-in’ likelihoods and latent models available
        - Temporal random effects, space-time interaction terms

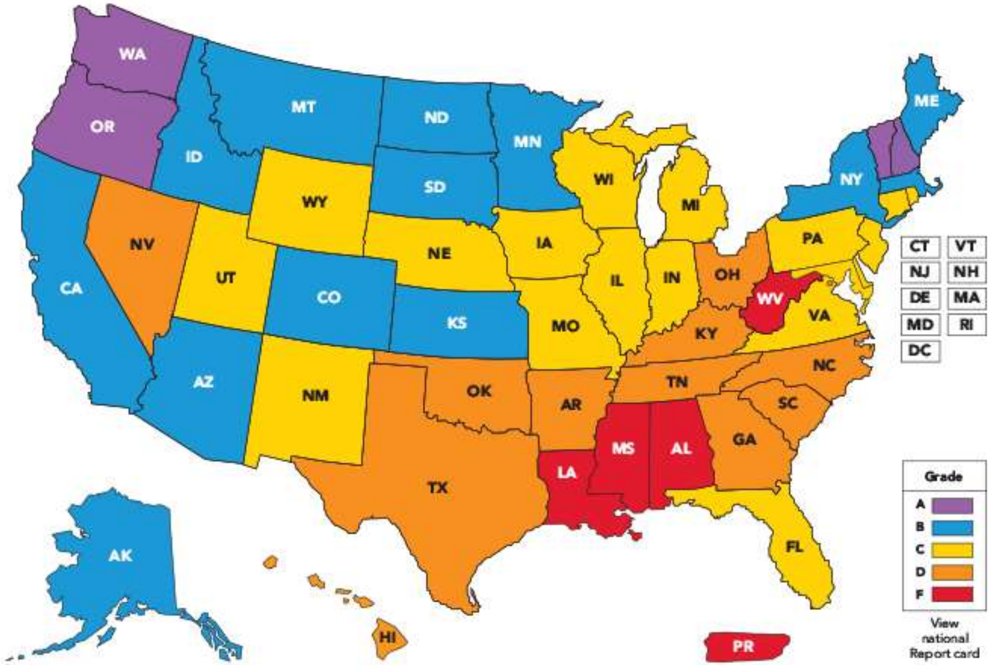
Other Approaches

- CARBayes in R
  - Intrinsic conditionally autoregressive models
  - Not as flexible as INLA
    - Gaussian, binomial, Poisson outcomes
  - MCMC simulations can be slow

- WinBUGS/OpenBUGS
  - Flexible
  - Slow, very computationally intensive
    - Can take ***weeks*** to run

Preterm birth rates

- Babies born too early have higher rates of death and other adverse health outcomes



https://www.marchofdimes.org/mission/prematurity-reportcard.aspx, SOURCE: National Vital Statistics System

INLA Model: Preterm birth rates

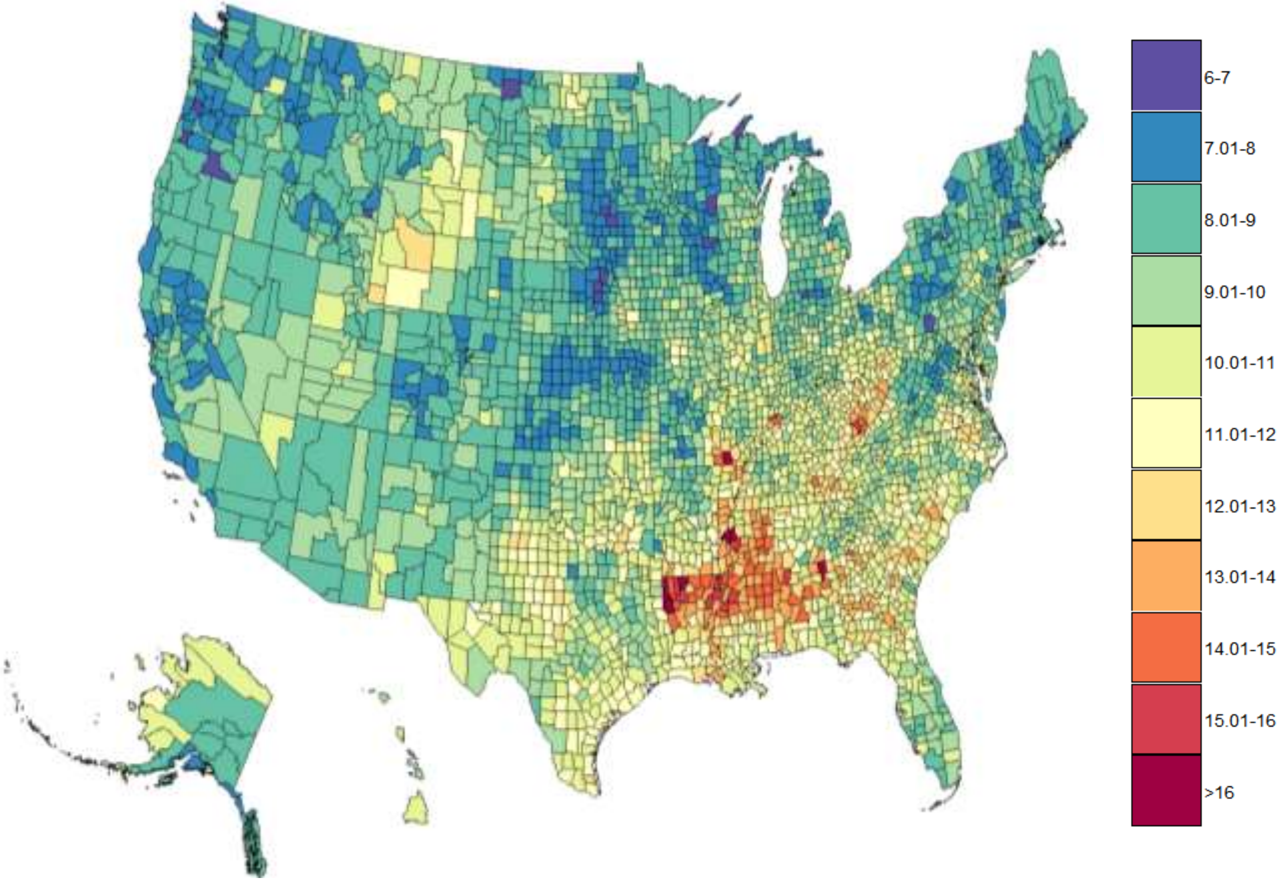
- Binomial models with spatially structured random effects:

$$Y_i \sim \text{Binomial}(N_i, p_i)$$
$$\text{logit}(p_i) = \alpha + u_i + v_i$$

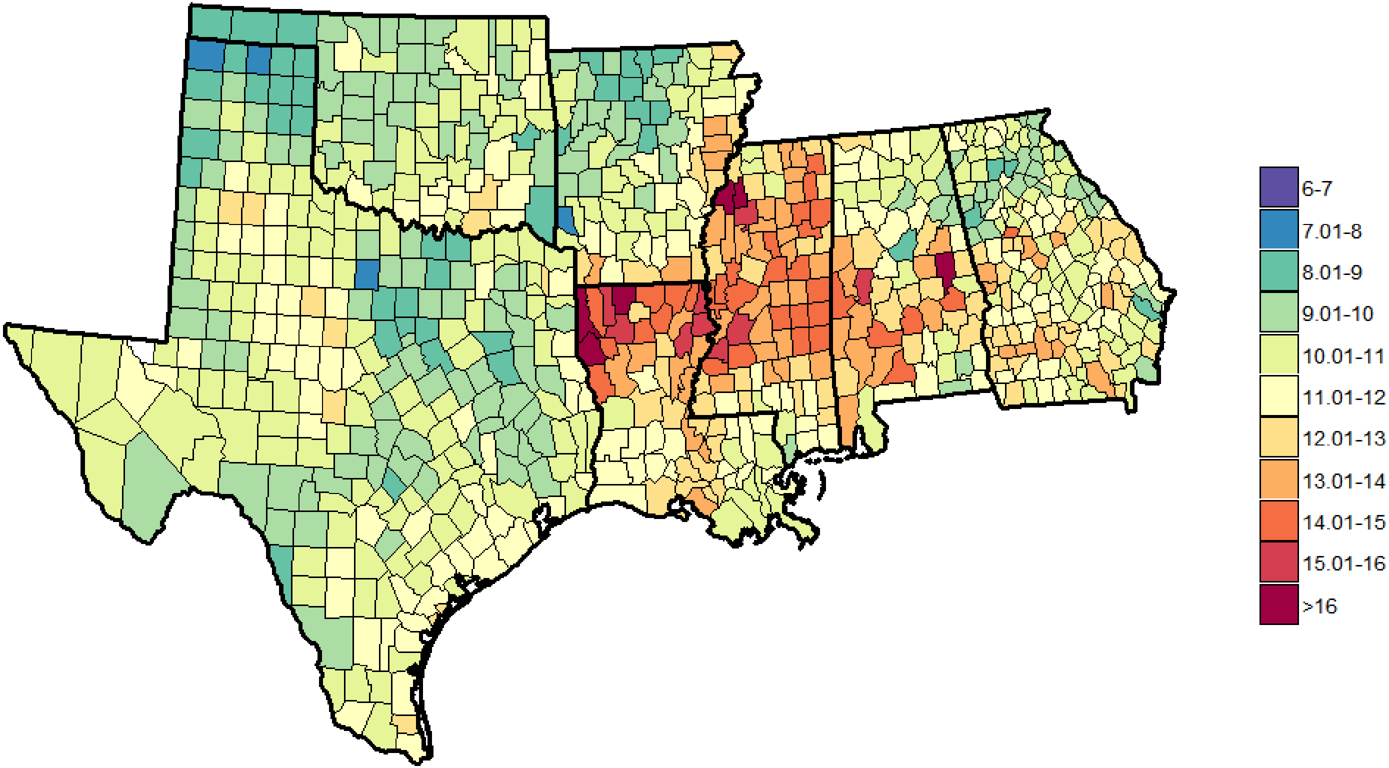
  - N_i = number of births in county i
  - α = intercept
  - u_i = spatially structured random effect
  - v_i = non-spatial random effect

- Compared results with:
  - Poisson, zero-inflated Poisson, zero-inflated binomial models (R-INLA)
  - Poisson and binomial models in CARBayes

County-level preterm birth rates, 2013-2015: INLA estimates

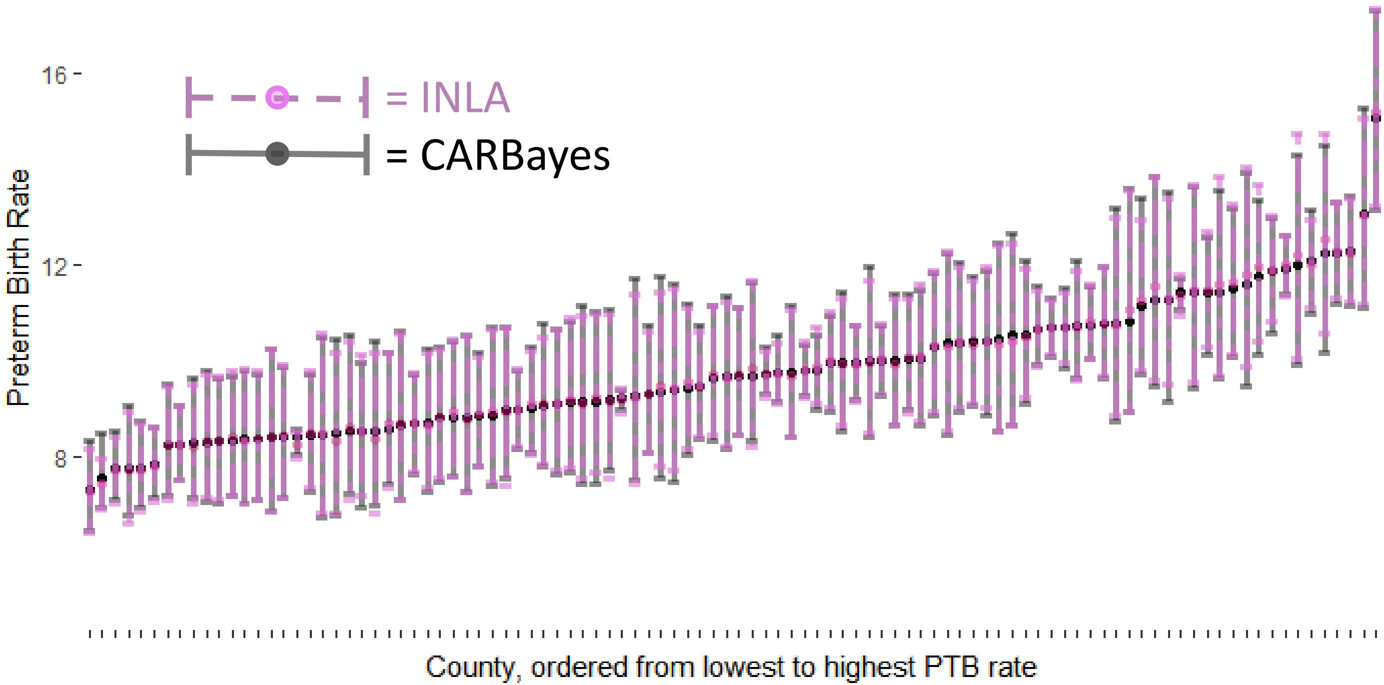


County-level preterm birth rates, 2013-2015: INLA estimates: selected states



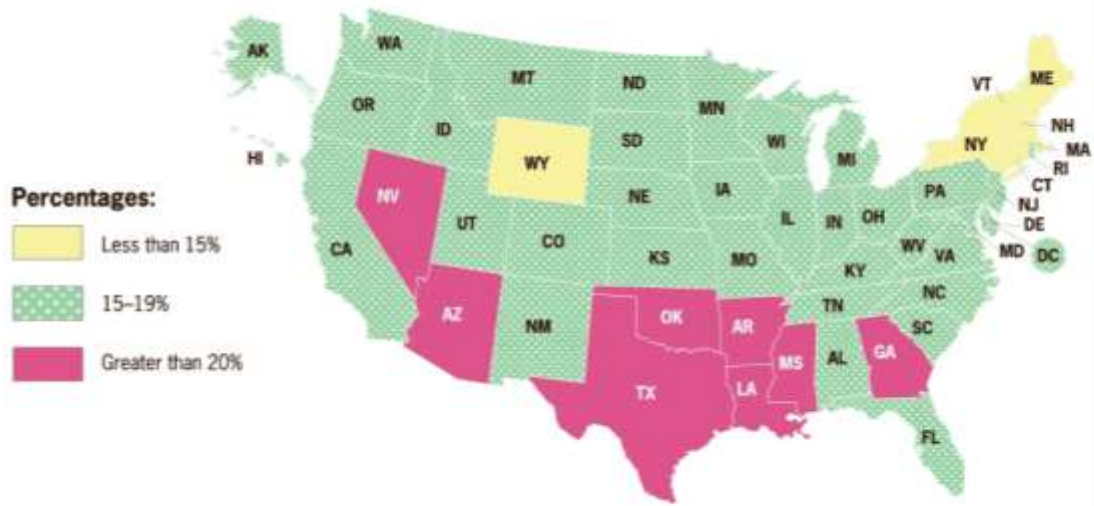
INLA vs. CARBayes

- Estimates and 95% credible intervals (CIs) very similar:



Second and higher order teen birth rates, 2007-2016

- Having more than one child as a teen is associated with negative health, emotional, social, and financial outcomes
  - Infants more likely to be born too early or too small
  - Limited educational and employment opportunities for the teen



https://www.cdc.gov/vitalsigns/teenpregnancy/index.html, National Vital Statistics System, teens, ages 15–19, 2010

INLA models: Second and higher order teen birth rates, 2007-2016

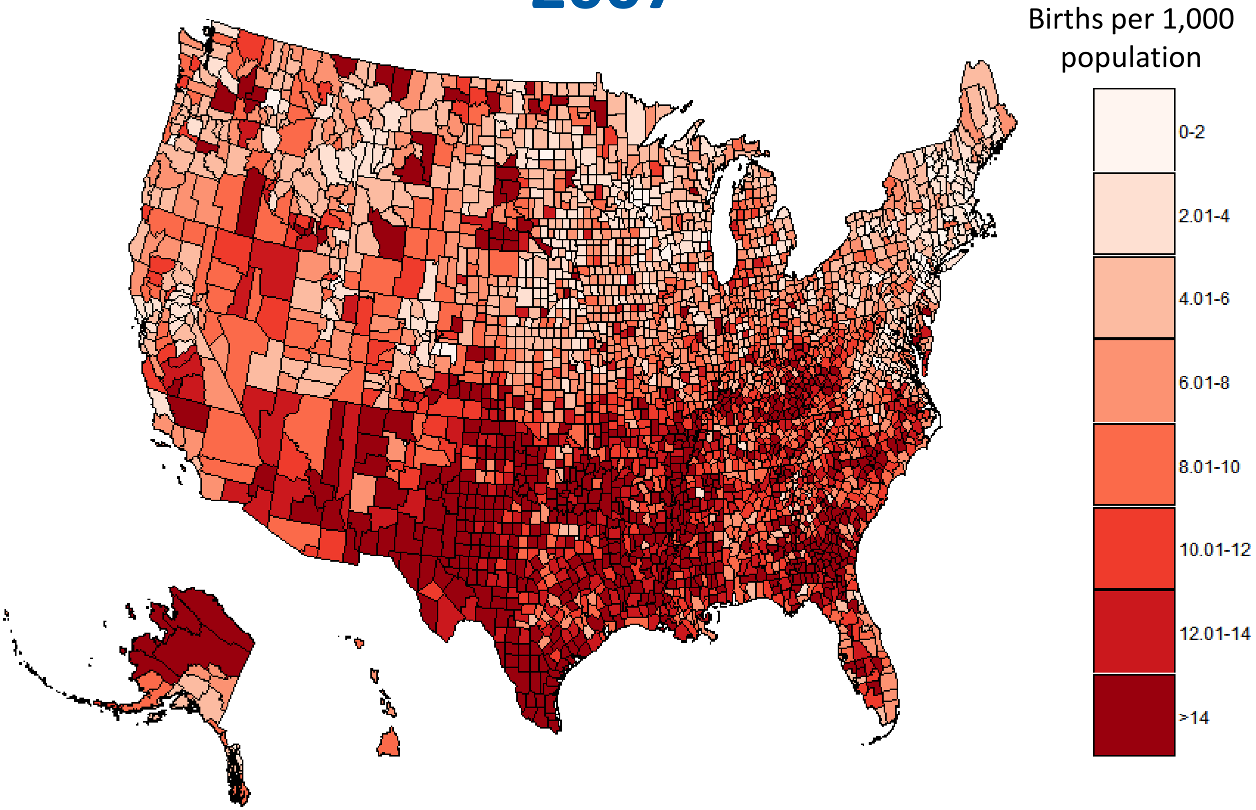
- Binomial space-time interaction models:

$$Y_{\mathrm{it}} \sim \mathrm{Binomial}(N_{\mathrm{it}}, p_{\mathrm{it}})$$

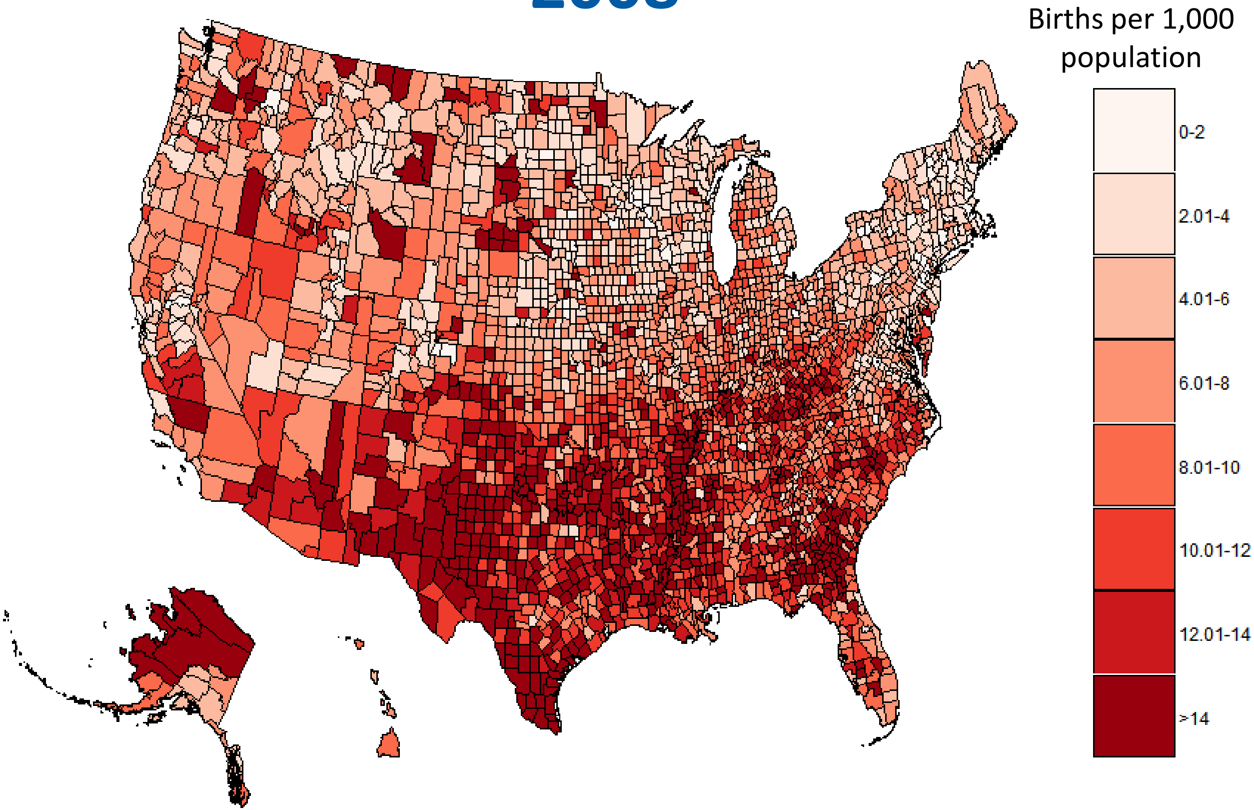
$$\mathrm{logit}(p_{\mathrm{it}}) = \alpha + \mathrm{A}_{\mathrm{i}} + \mathrm{B}_{\mathrm{t}} + \mathrm{C}_{\mathrm{it}}$$

  - N_{it} = number of births in county *i* at time *t*
  - p_{it} = probability of teen births in county *i* at time *t*
  - α = intercept
  - A_{i} = spatially structured random effect
  - B_{t} = time term
  - C_{it} = space-time interaction term

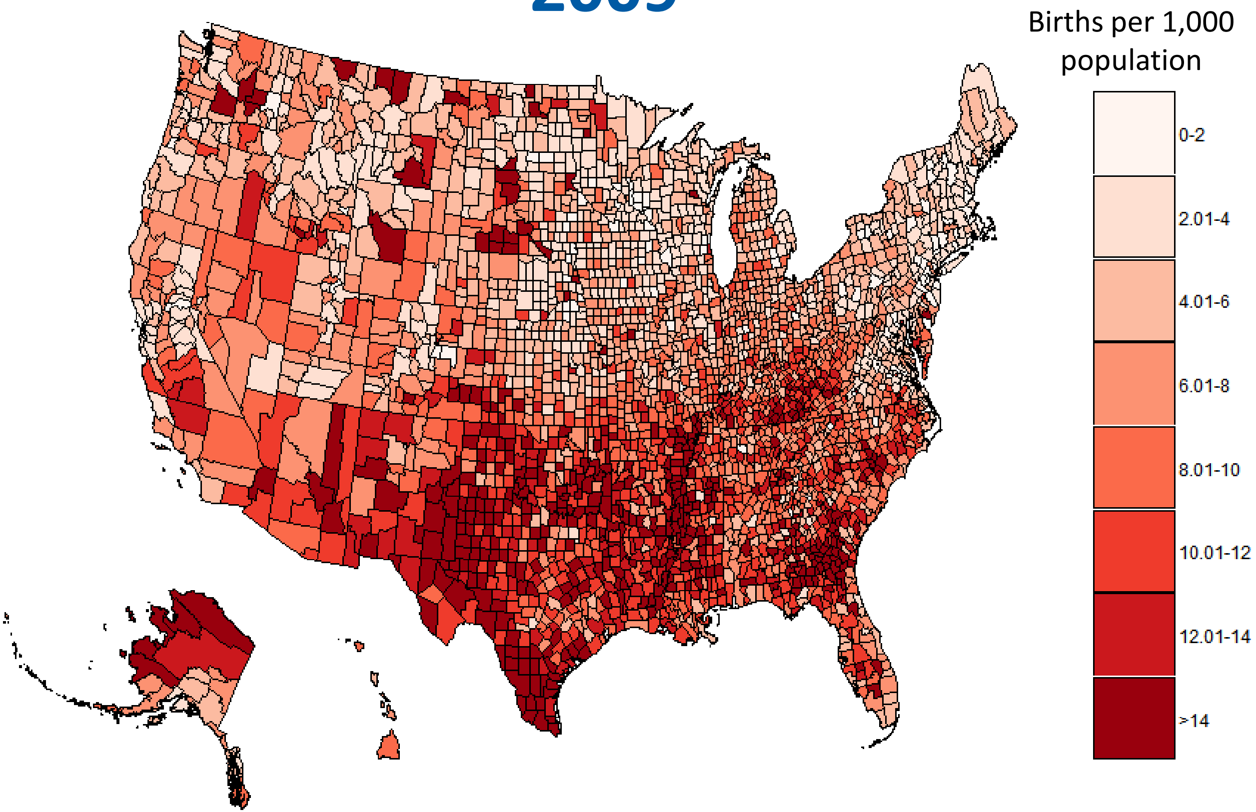
Second and higher order teen birth rates 2007



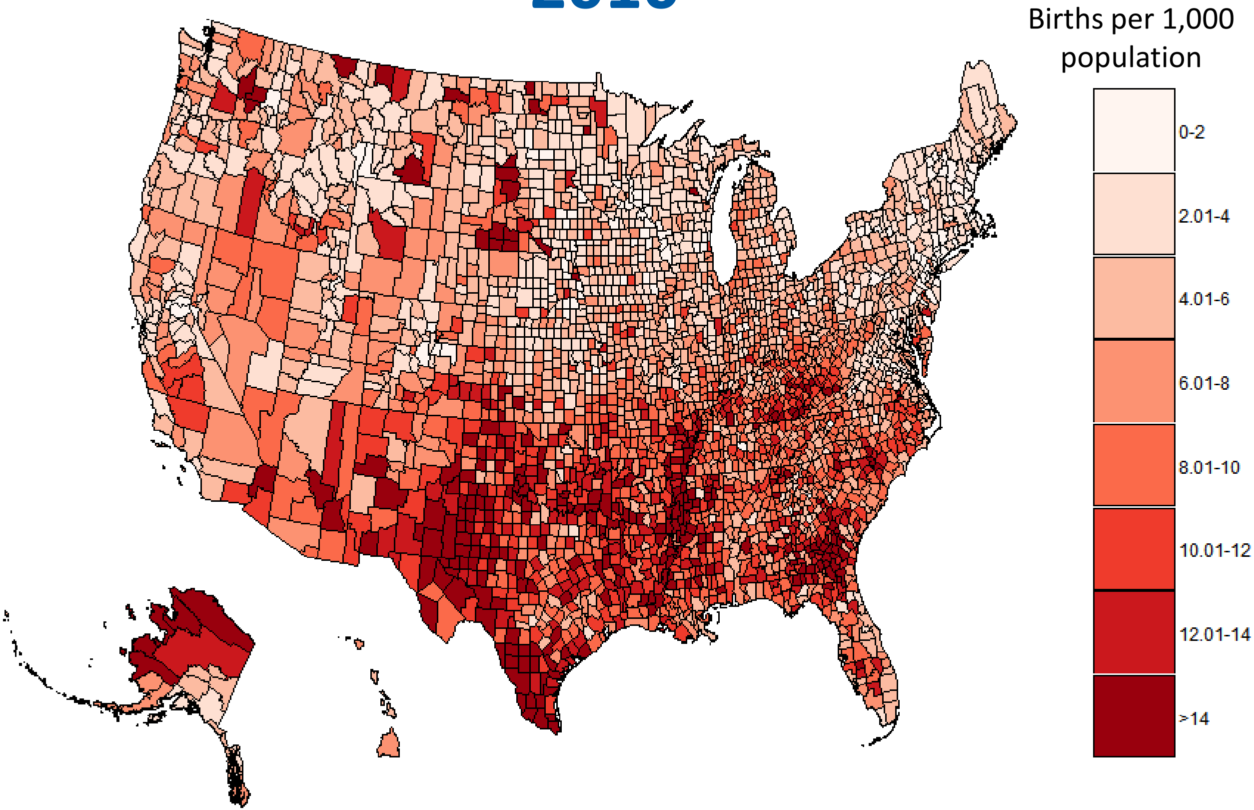
Second and higher order teen birth rates 2008



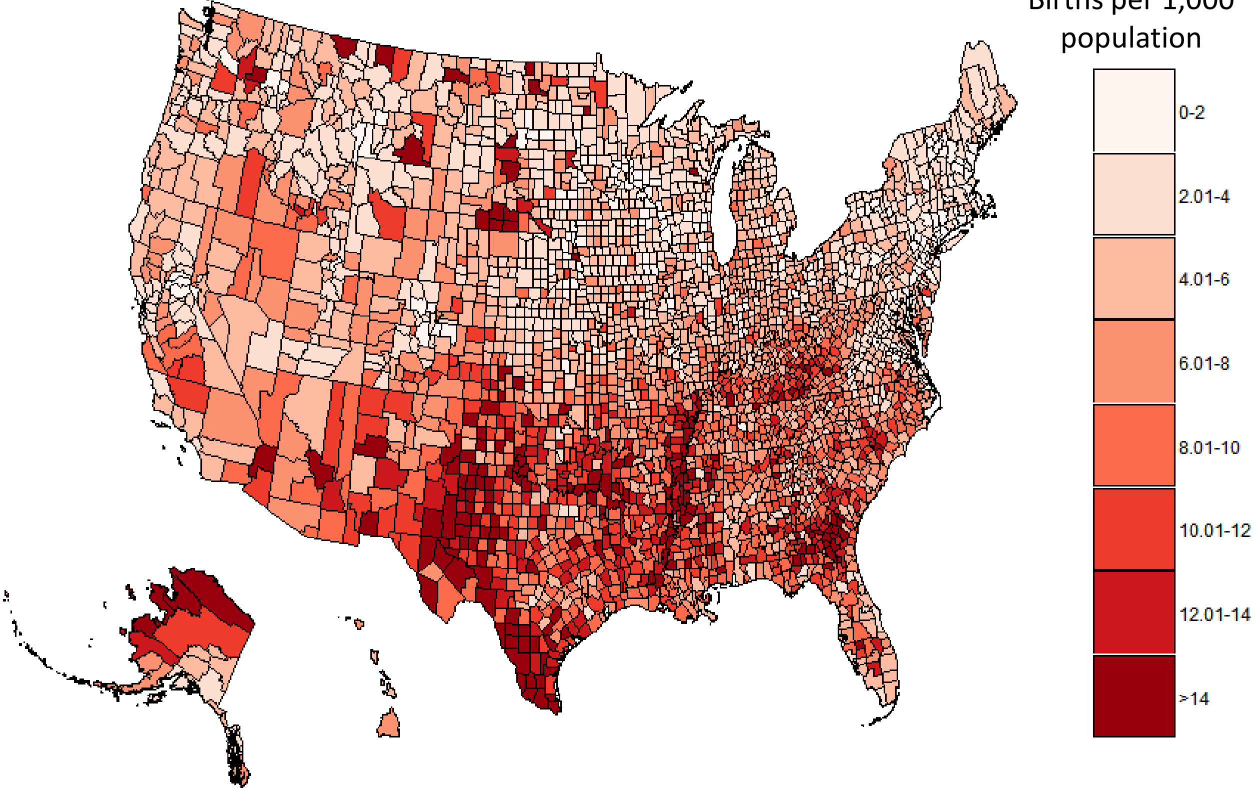
Second and higher order teen birth rates 2009



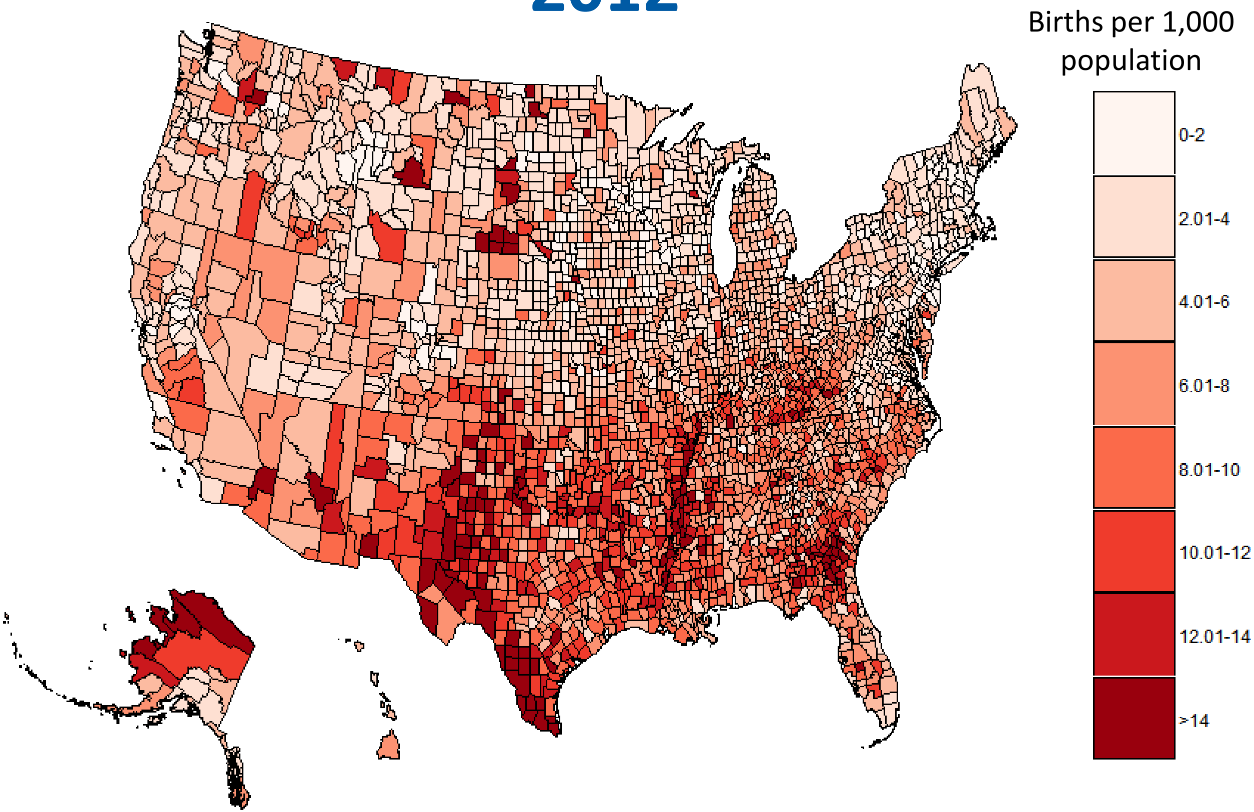
Second and higher order teen birth rates 2010



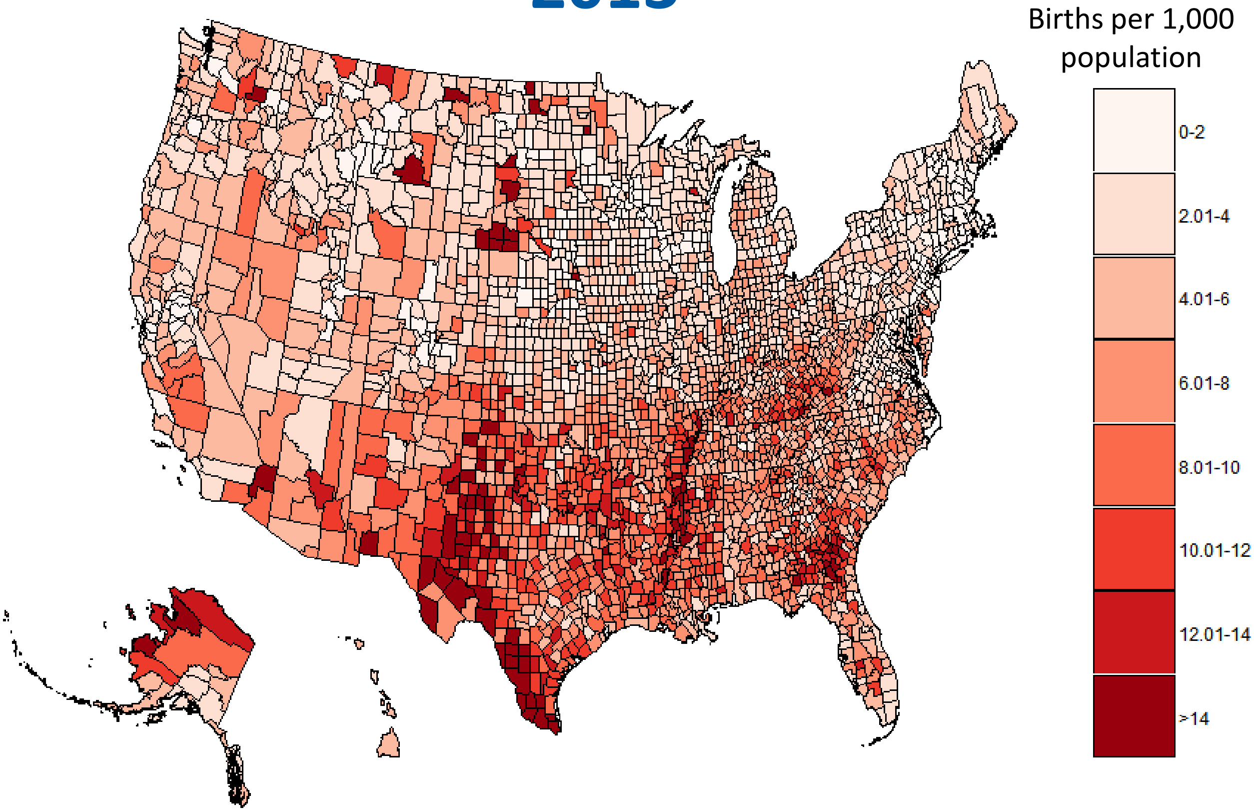
Second and higher order teen birth rates 2011



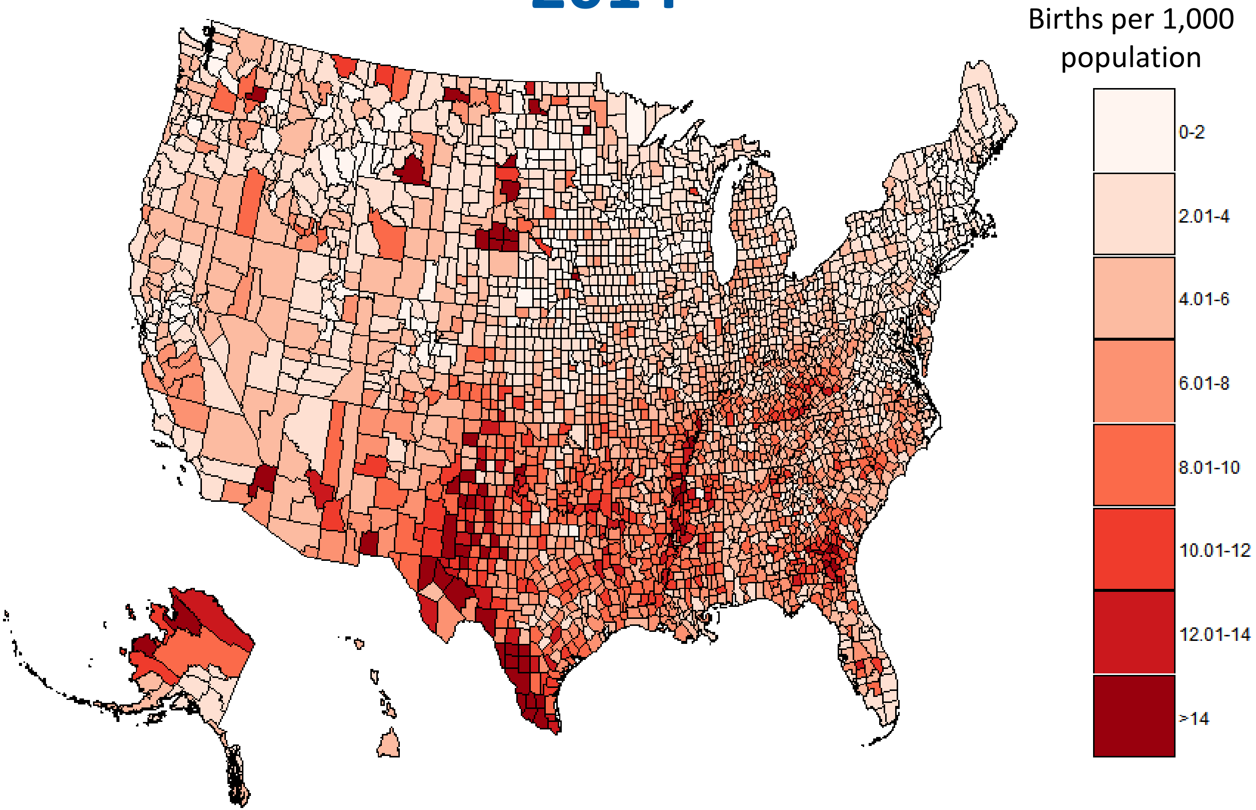
Second and higher order teen birth rates 2012



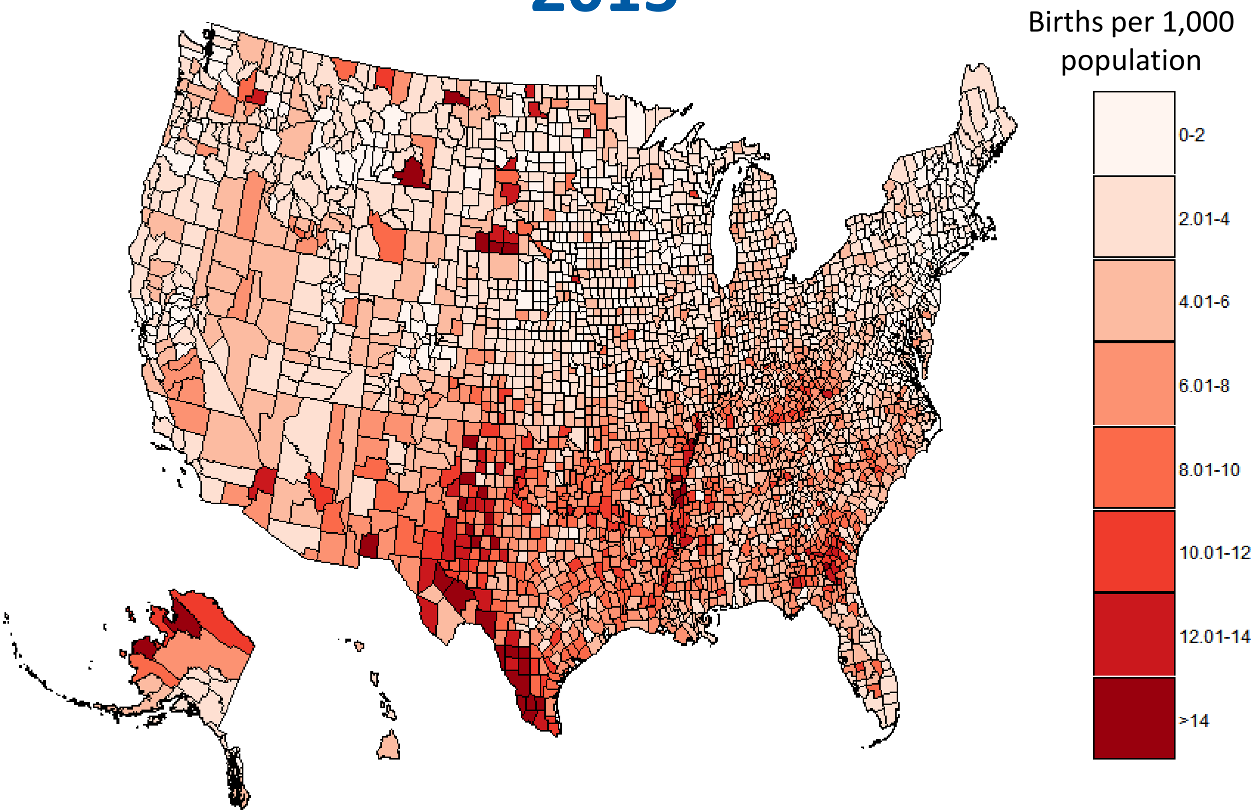
Second and higher order teen birth rates 2013



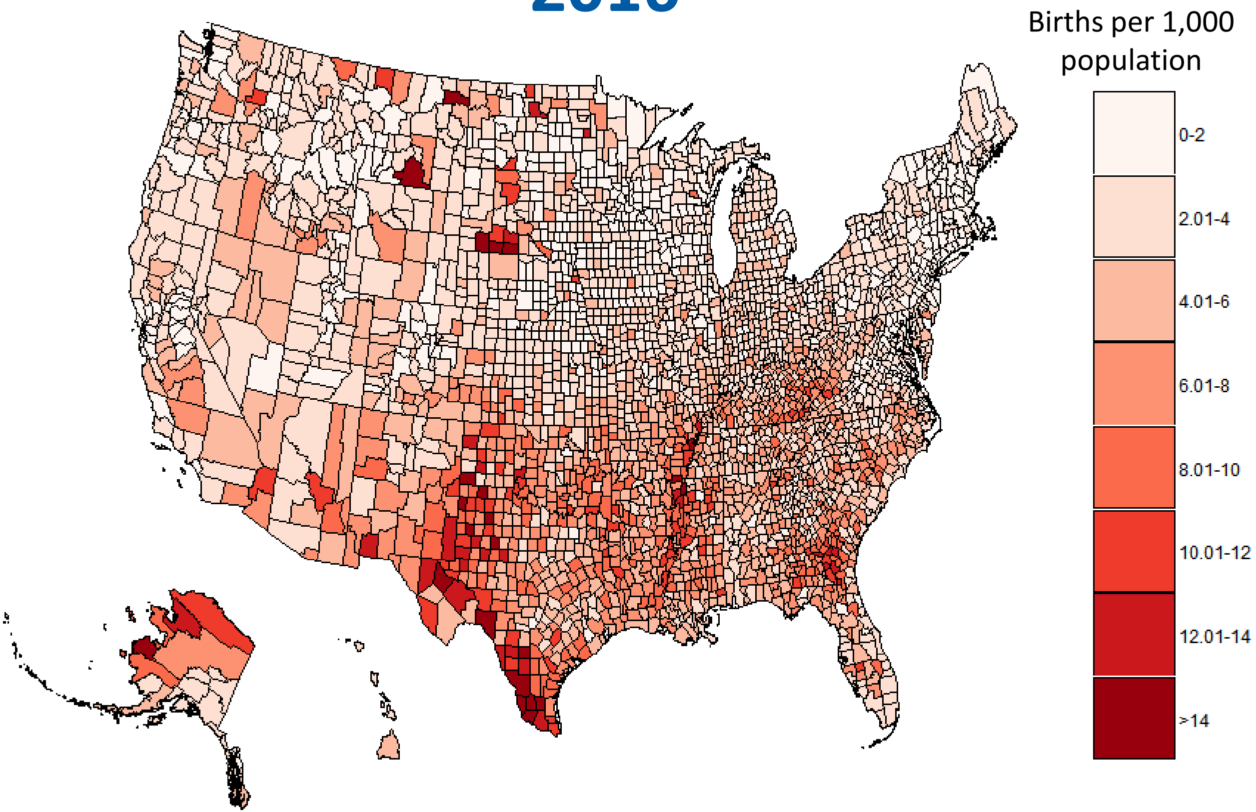
Second and higher order teen birth rates 2014



Second and higher order teen birth rates 2015



Second and higher order teen birth rates 2016



Infant Mortality Rates

- Considered a key marker of the overall health of a society
  - The United States has a higher infant mortality rate than similarly developed nations
- In 2015, 27 states met the Healthy People 2020 target of 6.0 infant deaths per 1,000 live births
  - Infant mortality rates higher in southern states

INLA models: Infant Mortality Rates, 2013-2015

- Zero-inflated Poisson models with spatially structured random effects

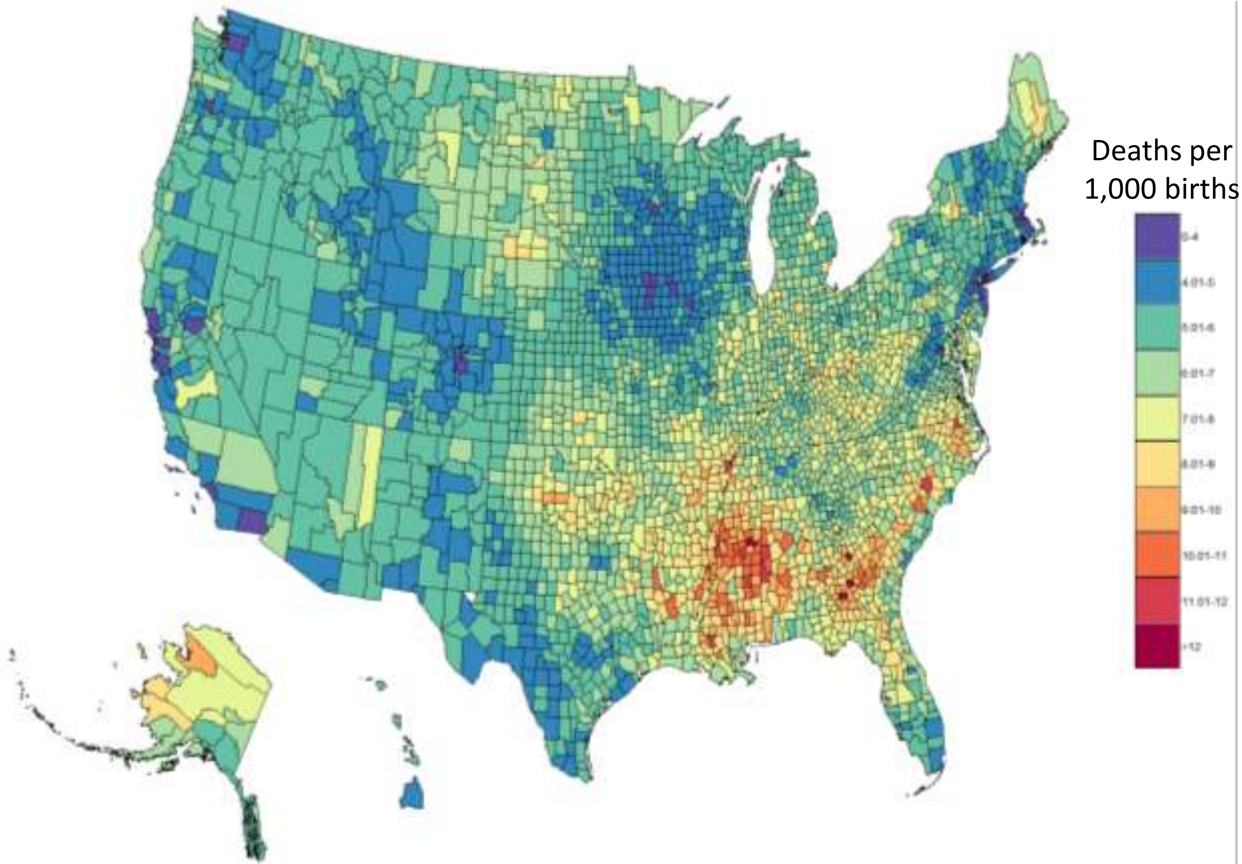
$$\text{Prob}(y| \ldots) \sim \begin{cases} 0, \text{ with probability } p \\ \text{Poisson}(y), \text{with probability } (1-p) \end{cases}$$

$$\log(y_{\text{i}}) = \alpha + u_{\text{i}} + v_{\text{i}} + \log(E_{\text{i}})$$

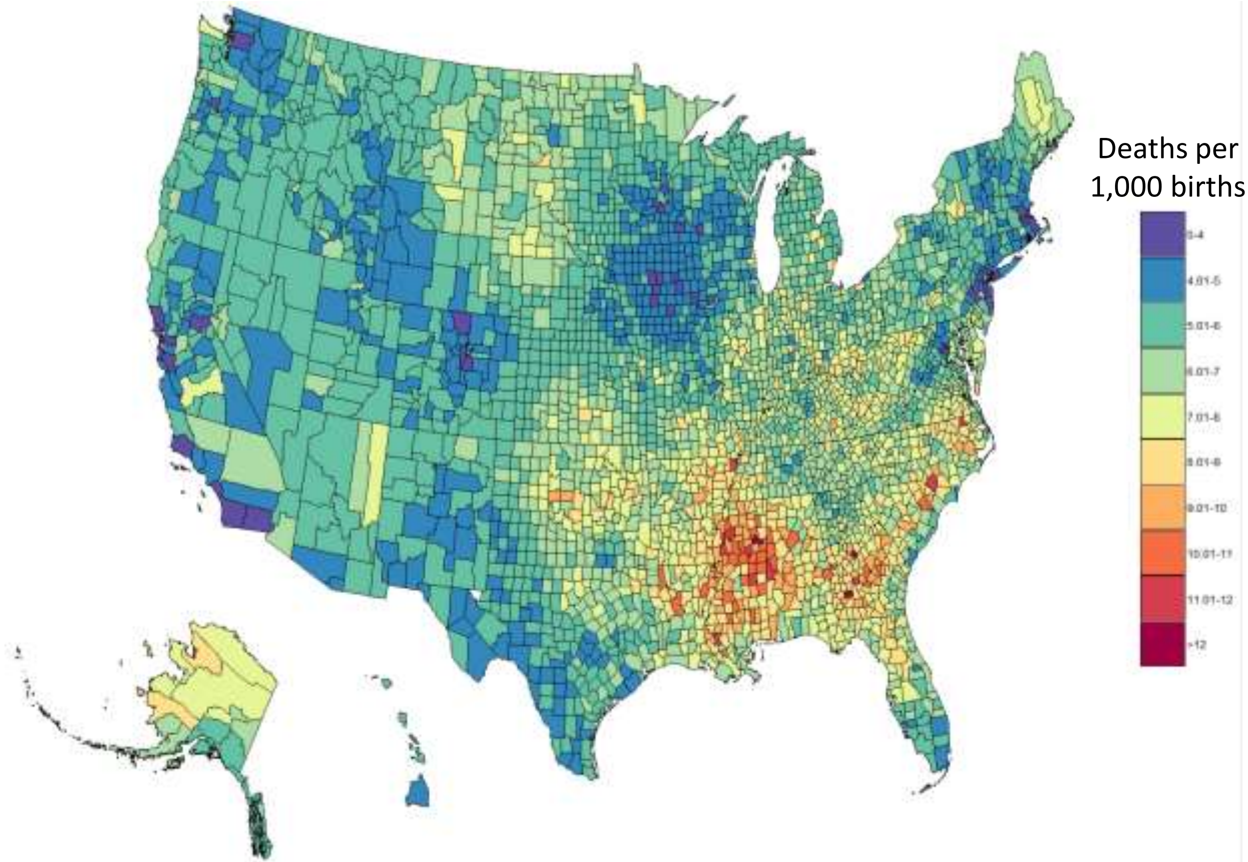
  - E_i = exposure, number of births in county i
  - α = intercept
  - u_{i} = spatially structured random effect
  - v_{i} = non-spatial random effect

- Compared results with:
  - Poisson, binomial, zero-inflated binomial models (R-INLA)
  - Poisson and binomial models in CARBayes

Infant deaths per 1,000 live births, 2013-2015: INLA estimates



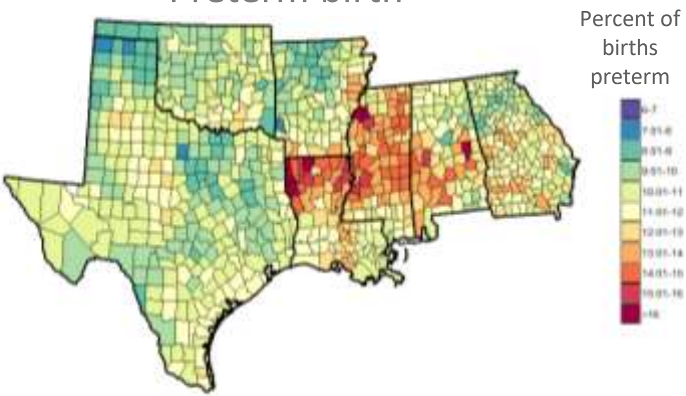
Infant deaths per 1,000 live births, 2013-2015: CARBayes estimates



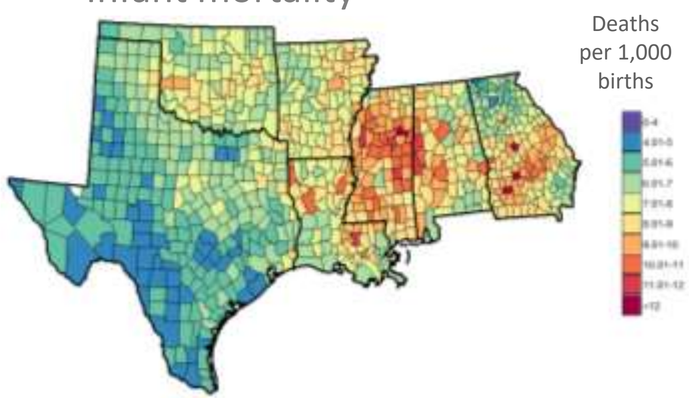
Discussion

– Birth or death rates at the county level are often unstable, suppressed for small areas
– Aggregating over several years or larger geographic regions can mask patterns and trends
  - Variation within states or over time
  - Areas of high or low values that cross state boundaries

Preterm birth



Infant mortality



Limitations and Strengths

- Model-based estimates might smooth away important effects
- People trust direct estimates (*real data*) more
  - "Black box" models, assumptions

- Various model-based approaches produce rather consistent results
  - For a variety of birth and death outcomes examined
    - INLA, CARBayes, WinBUGS/OpenBUGS
    - Different likelihoods and models with/without covariates
  - The overall patterns are very similar

Conclusions

- Model-based approaches can be used to generate county-level estimates of birth and death rates
  - Examine variation across the entire U.S.
  - Pick up on important spatial or temporal patterns that might be masked by state estimates or other groupings (urban/rural)
  - Provide information relevant to public health efforts at the state or local level
  - Shed light on risk/protective factors associated with population health outcomes

Questions?

Lauren Rossen
LRossen@cdc.gov

Division of Vital Statistics
National Center for Health Statistics

Diba Khan
ild1@cdc.gov

Division of Research Methodology
National Center for Health Statistics

INLA Models

- **Preterm birth**

```
> numerator ~ 1 + f(region, model="bym", graph="map")
> inla(formula,family="binomial",Ntrials=denominator, data=data,
    control.compute=list(dic=TRUE, cpo=TRUE, waic=T))
```

- **Teen birth rates**

```
> numerator ~ 1 + year + f(region, model="bym", graph="map") +
    f(interaction, model="rw1")
> inla(formula,family="binomial",Ntrials=denominator, data=data,
    control.compute=list(dic=TRUE, cpo=TRUE, waic=T))
```

- **Infant mortality**

```
> numerator ~ 1 + f(region, model="bym", graph="map")
> inla(formula, family="zeroinflatedpoisson1", E=denominator,
    data=data, control.compute=list(dic=TRUE, cpo=TRUE, waic=T))
```

Helpful References

- http://www.r-inla.org/
- Bivand R, Sha Z, Osland L, Thorsen IS. A comparison of estimation methods for multilevel models of spatially structured data. *Spatial Statistics* 2017;21:440-459.
- Blangiardo M, Cameletti M, Baio G, Rue, H. Spatial and spatio-temporal models with R-INLA. *Spatial & Spatio-temporal Epidemiology*. 2013;4:33-49.
- Carlin BP, Louis TA. 2009. Bayesian Methods for Data Analysis. New York: Chapman and Hall.
- Lawson A. 2013. Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. New York: Chapman and Hall.
- Lawson A, Biggeri AB, Boehning D, Lesaffre E, Viel JF, Clark A, Schlattmann P, Divini F. Disease mapping models: an empirical evaluation. *Statistics in Medicine* 2000;19:2217-2241.