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

¹Division of Vital Statistics, National Center for Health Statistics
²Division 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

**DISCLAIMER:** The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Center for Health Statistics or the Centers for Disease Control and Prevention.
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

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

[Map showing infant mortality rates across the United States]
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

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 \( i \)
  - \( \alpha \): 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:

![Graph showing preterm birth rates across counties, with INLA and CARBayes methods compared.](image)
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

Second and higher order teen birth rates, 2007-2016

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

- Binomial space-time interaction models:
  \[ Y_{it} \sim \text{Binomial}(N_{it}, p_{it}) \]
  \[ \logit(p_{it}) = \alpha + A_i + B_t + C_{it} \]
  - \( N_{it} \) = number of births in county \( i \) at time \( t \)
  - \( p_{it} \) = probability of teen births in county \( i \) at time \( t \)
  - \( \alpha \) = 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

Births per 1,000 population

- 0-2
- 2.01-4
- 4.01-6
- 6.01-8
- 8.01-10
- 10.01-12
- 12.01-14
- >14
Second and higher order teen birth rates
2008
Second and higher order teen birth rates 2009

Births per 1,000 population
Second and higher order teen birth rates
2010

![Map of the United States showing birth rates per 1,000 population for second and higher order teen births in 2010. The map uses a color scale ranging from light pink (0-2 births per 1,000 population) to dark red (12.01-14 births per 1,000 population).](image)
Second and higher order teen birth rates 2011

Births per 1,000 population
Second and higher order teen birth rates 2012

Births per 1,000 population
Second and higher order teen birth rates
2013
Second and higher order teen birth rates 2014

Births per 1,000 population

- 0-2
- 2.01-4
- 4.01-6
- 6.01-8
- 8.01-10
- 10.01-12
- 12.01-14
- >14
Second and higher order teen birth rates
2015
Second and higher order teen birth rates
2016

Births per 1,000 population

[Map of the United States showing color-coded birth rates]
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

- Zero-inflated Poisson models with spatially structured random effects

\[
\text{Prob}(y | \ldots) \sim \{ \begin{array}{ll}
0, & \text{with probability } p \\
\text{Poisson}(y), & \text{with probability } (1-p)
\end{array}
\]

\[
\log(y_i) = \alpha + u_i + v_i + \log(E_i)
\]

- \(E_i\) = exposure, number of births in county \(i\)
- \(\alpha\) = 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](image1)

![Infant mortality](image2)
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
  \[ \text{numerator} \sim 1 + f(\text{region}, \text{model}="bym", \text{graph}="map") \]
  \[ \text{inla(formula,family}="\text{binomial}",\text{Ntrials}=\text{denominator}, \text{data}=\text{data}, \text{control.compute}=\text{list}(\text{dic=}TRUE, \text{cpo=}TRUE, \text{waic=}T)) \]

- Teen birth rates
  \[ \text{numerator} \sim 1 + \text{year} + f(\text{region}, \text{model}="\text{bym}", \text{graph}="map") + f(\text{interaction}, \text{model}="\text{rwl}") \]
  \[ \text{inla(formula,family}="\text{binomial}",\text{Ntrials}=\text{denominator}, \text{data}=\text{data}, \text{control.compute}=\text{list}(\text{dic=}TRUE, \text{cpo=}TRUE, \text{waic=}T)) \]

- Infant mortality
  \[ \text{numerator} \sim 1 + f(\text{region}, \text{model}="\text{bym}", \text{graph}="map") \]
  \[ \text{inla(formula, family}="\text{zeroinflatedpoisson1}", \text{E}=\text{denominator}, \text{data}=\text{data}, \text{control.compute}=\text{list}(\text{dic=}TRUE, \text{cpo=}TRUE, \text{waic=}T)) \]
Helpful References

- [http://www.r-inla.org/](http://www.r-inla.org/)