Nebraska Center for Rural Health Research

University of Nebraska Medical Center

A Comparison of Spatial Analysis Techniques:

A Case Study Using Nebraska Pediatric Cancer Data

Presenter
Michael Shambaugh-Miller, Ph.D.

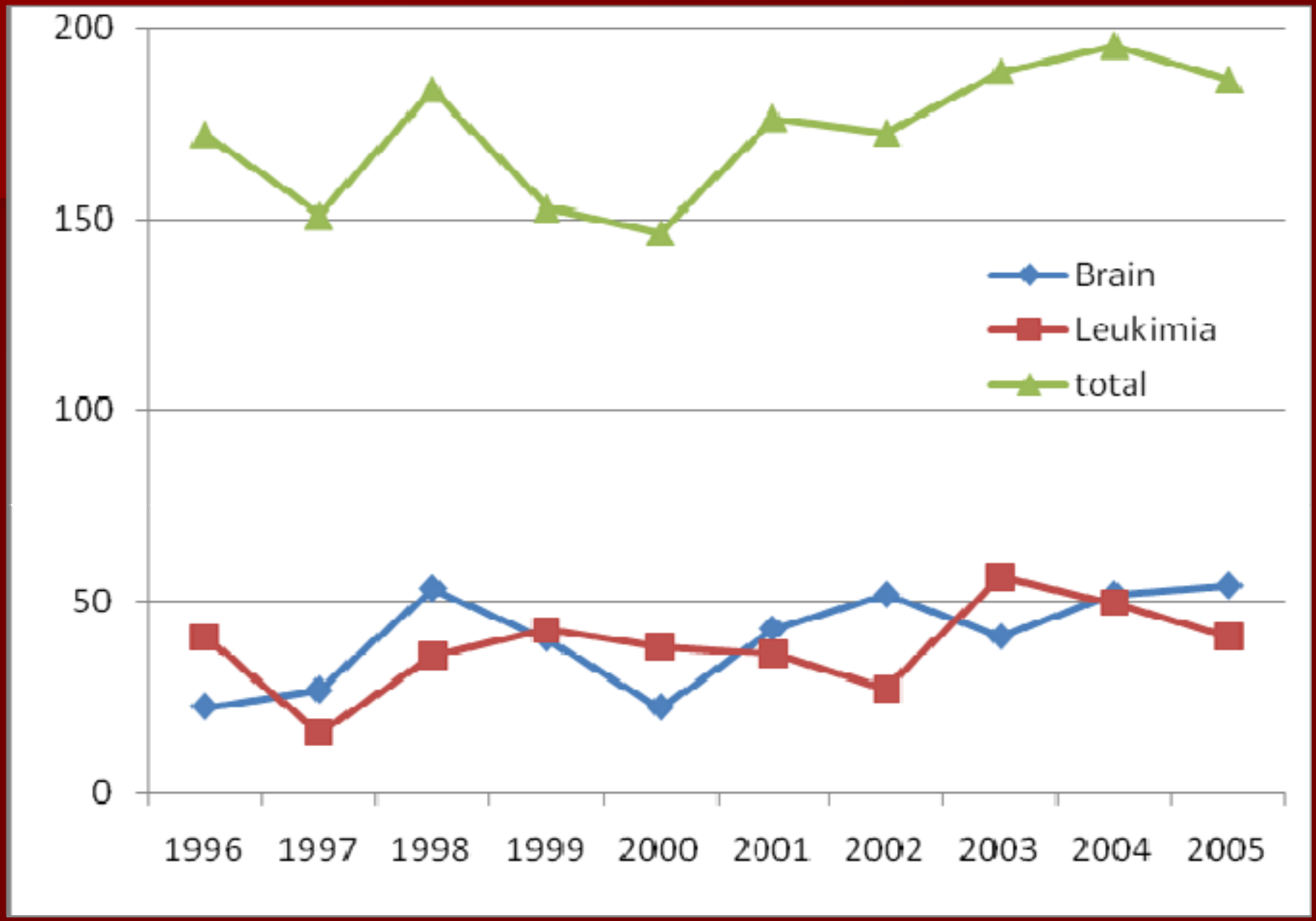
Co-Authors
Jane Meza, Ph.D. Dept. of Biostatistics
Ge Lin, Ph.D. Dept. of Health Services and Administration
Nicole VanOsdel, M.A. Dept. of Health Services and Administration
Kate Hanson, M.A. Dept. of Biostatistics

College of Public Health
University of Nebraska Medical Center

Study Rationale

- Even though childhood cancer is rare, it is a major cause of children's death in the United States.
- In this report, we define children as 17 or younger, and there were 445,087 of them in Nebraska in 2005 with 87 cancer cases.
- During the 10 year between 1996 and 2005, there were 773 childhood cancer cases, and the average incidence rate is 172.6 per million (Figure 1).

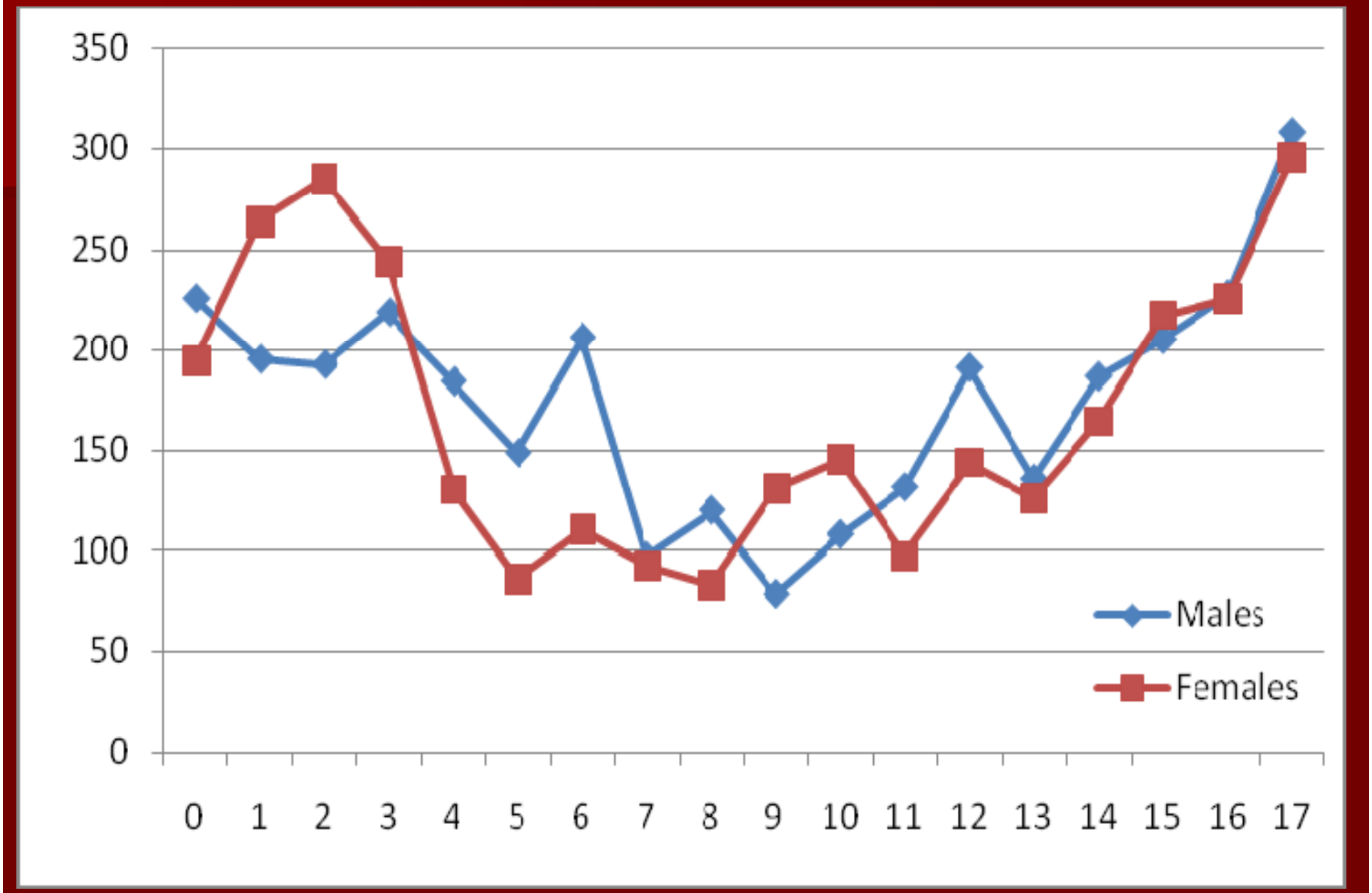
Figure 1. Childhood cancer incidence rates for selected sties



Study Rationale

- Although not exactly comparable due to age difference, this rate is slightly higher than the US average of about 150 for those age 20 or younger.
- The cancer rates by race and sex or age are all similar to the general characteristics of the US children.
  - Blacks (12.8% vs 16% for Whites)
  - females (16.8% vs 17.5% for males) have lower rates
  - both younger and older children have higher rates (Figure 2).

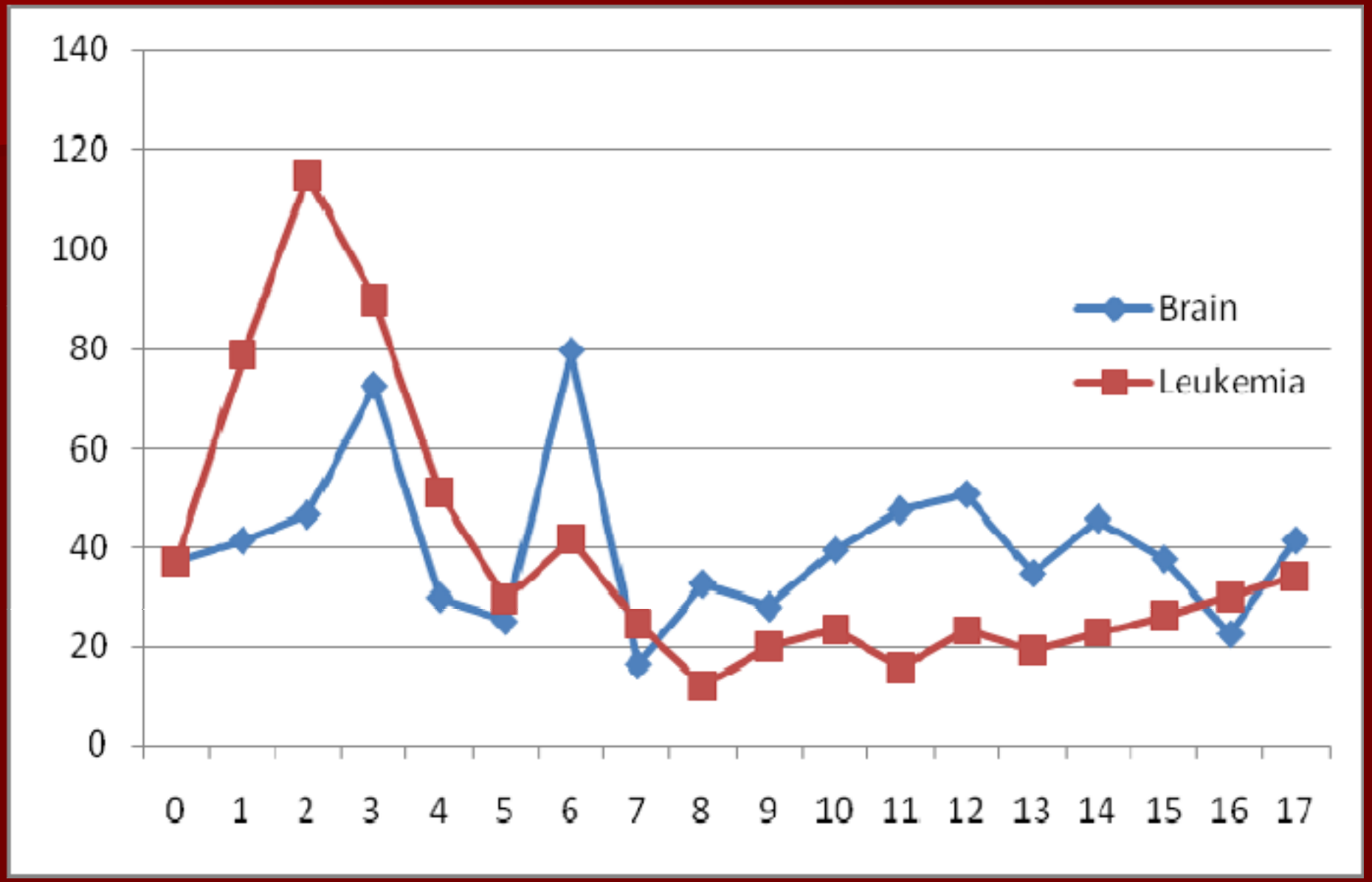
Figure 2. Childhood cancers by age and sex: 1996-2005



Study Rationale

- The limited case numbers prohibit most site specific analyses, but two major sites, namely brain and leukemia account for 23.4 and 22.0 percents, respectively, of all children cancers (Figure 3).

Figure 3. Childhood cancers by age for selected sties: 1996-2005



Study Rationale

- Leukemia incidence rate for Nebraska children were about 38 percent, and it is comparable to the national average of 37 per million.
- It seems that Nebraska children had higher brain cancer incidence rate (40.4 per million) than the nation (25 per million).
- Again the age specific trends are similar to the national trends, and the Nebraska trends tend to vary more than the national trend due to small number of cases, and the slight increase in brain cancer in Figure 1, therefore, may not necessarily suggest a trend.

Study Methodology

- We used a model based estimation approach to develop a *small-area estimation model* to be used to obtain estimates of pediatric cancer rates for each of Nebraska's 93 counties.
- The registry data were used to construct preliminary rates by county. For counties with zero or very few observed cases, the estimates based on county-specific samples are likely to be imprecise. On the other hand, estimates from counties with large observed frequencies will have greater precision.

Study Methodology

- The preliminary rates by county are based on the standardized mortality ratio (SMR).
- The SMR, an estimate of relative risk, is the ratio of the number of observed cases (*Oi*) to the number of expected cases (*Ei*) within a geographic unit *i* (i.e., the odds of being in the disease group rather than in the background group).
- In this part of study, we use county as the unit of geographic analysis.
- The observed number of cases for each county was obtained for the registry data.
- The expected number of cases for each county was based on the state wide rate using the Nebraska registry data.

Study Methodology

- While the SMR is a commonly used measure of relative risk, it has limitations.
- The variance of the SMR is large for regions with small expected values (i.e., counties with small population size).
- In addition, the SMR is always 0 when there are no observed cases, regardless of the population size (Meza, 2003).

Study Methodology

- Due to the difficulties in interpreting the SMR, statistical methods which borrow information from other geographic areas and/or covariates have been developed.
- These statistical methods include empirical Bayes and Bayesian approaches, which estimate county level rates by “borrowing” information from the full data set, reducing the variance of the estimates and “shrinking” county level estimates toward the overall mean SMR for the collection of regions.
- Estimates from areas with unreliable SMRs due to small population size will rely more on the overall mean than areas with larger populations (Meza, 2003).
- SMRs were determined for each county, followed by the application of a Bayesian spatial conditional autoregressive model to analyze the geographical distribution of NHL in relation to atrazine concentration and establish relative risk estimates.

- It is well known that disease incidence rates vary geographically. Many statistical models that account for correlated relative focus on geographical dependent correlations and can be extended to examine the covariate associations.
- A random-effects Poisson model to account for spatial conditional modeling using the intrinsic conditional autoregressive (CAR) model of Besag, York and Mollie (1991) was used to smooth the county level SMRs toward the mean of the adjacent neighboring counties. The model is given by:

$$Oi \sim Poisson\ (\mu I)$$
$$\log(\mu I) = \log(EI) + \alpha 0 + bi$$

Results

- The median age at diagnosis was 9.8 years (minimum 0 months, maximum 17.9 years). The 25th percentile was 3.4 years and the 75th percentile was 15.0 years. The distribution of age at diagnosis is displayed in Figure 4. Table 1 displays the distribution of sex, race and year of diagnosis.

Figure 4. Histogram of age (years) at diagnosis, NE pediatric cancers 1996-2005

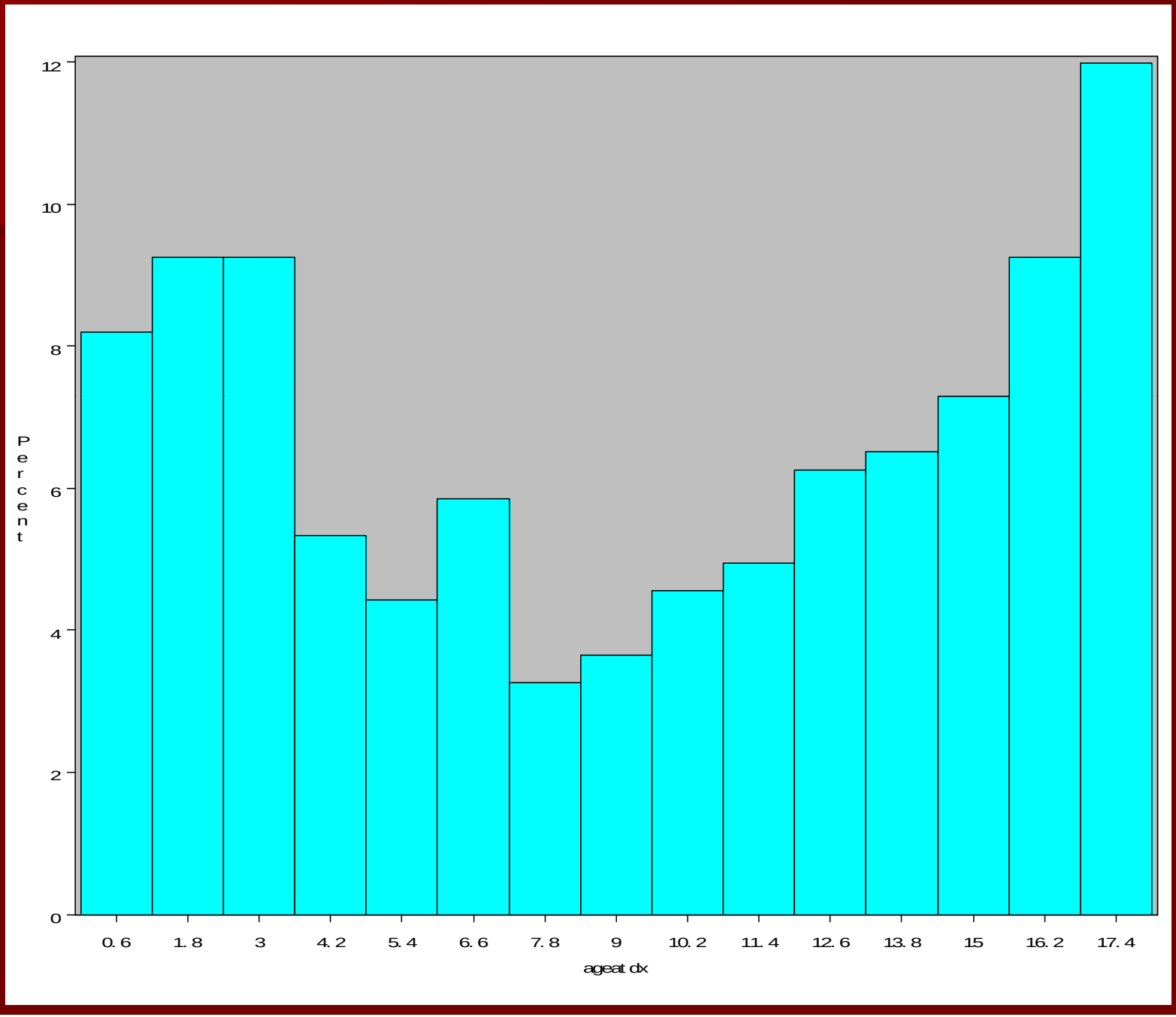


Table 1. Distribution of sex, race and year, NE pediatric cancers 1996-2005

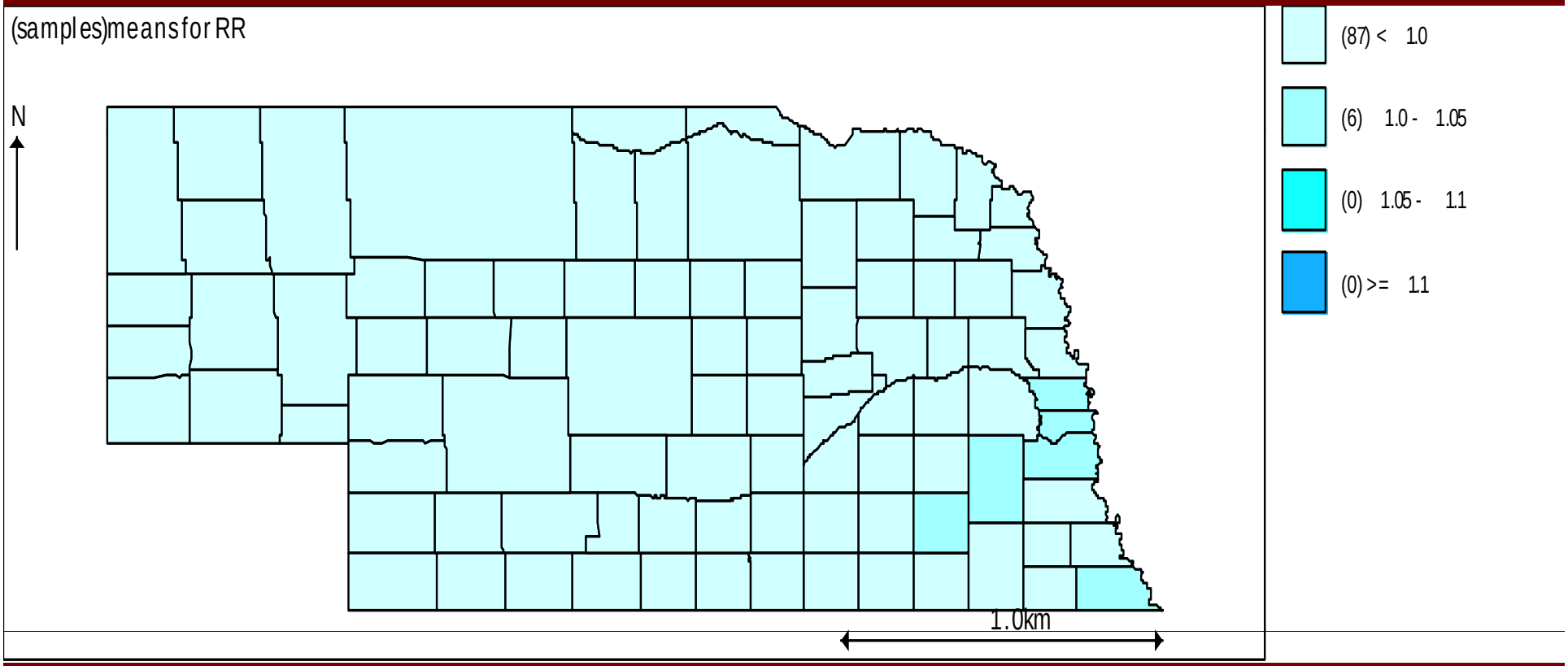
SEX				
SEX	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
Male	400	52.08	400	52.08
Female	368	47.92	768	100.00

Race	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
White	684	89.06	684	89.06
Black	36	4.69	720	93.75
Other	24	3.13	744	96.88
Unknown	24	3.13	768	100.00

County level analysis using supraregional rate based Nebraska Cancer Registry Data

- Using the Nebraska Cancer Registry to determine the supraregional rate, there was little variation in the distribution of pediatric cancer cases by county (Figure 5).

Figure 5. CAR smoothed SMRs,
Supraregional rate = Nebraska rate using Nebraska Cancer Registry Data



County level analysis using supraregional rate based regional and national rates

- A similar analysis was conducted using supraregional rates of 165.92 per million for the nation and 165.50 per million for the Midwest based on results of analysis of cancer registry data from 2001-2003 for children and adolescents age 0-19 in Li et al (2003).
- Results using these rates also revealed little variation in the distribution of cancer cases by county in Nebraska.

County level analysis using supraregional SaTScan

- County level analysis using SaTScan also revealed little variation in the distribution of pediatric cancer cases by county.

Health planning region analysis using SaTScan

- The small amount of variation in pediatric cancer cases by county may be due to true homogeneity of the pediatric cancer rates, or may be a result of oversmoothing due to the small number of observed cases. Therefore, an additional analysis was conducted using the 6 health planning regions. Analysis at the health planning region level also did not reveal any significant variation in pediatric cancer cases.

Conclusions and recommendations

- Little variation in pediatric cancer cases by county was observed. This may be due to true homogeneity of the pediatric cancer rates, or may be a result of oversmoothing due to the small number of observed cases. Analysis at the local public health department region also did not reveal any significant variation in pediatric cancer cases.

Conclusions and recommendations

- According to Li et al (2008), nationally, boys had a higher rate of cancer compared with girls; children aged 0-14 had a lower rate than adolescents age 15-19; and whites had a higher rate of cancer compared with other races. Additional analyses should be conducted to explore whether these associations hold in Nebraska.

Continued Studies

- Rurality
- Socioeconmic impact
- Other child health issues
  - Dentistry
  - Prenatal access
  - Postnatal access
  - Health education

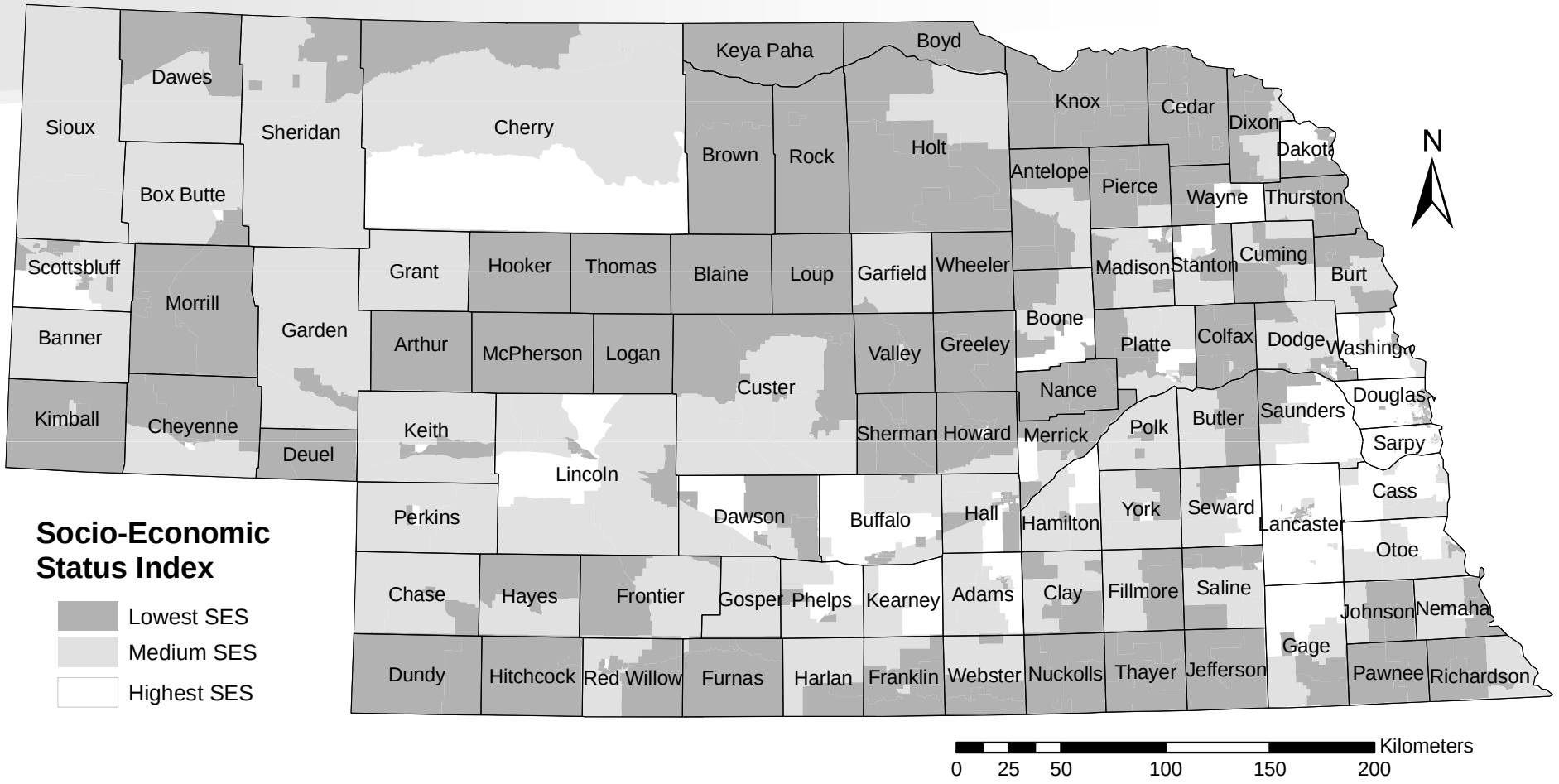
Pediatric Cancers Analysis by Rurality

Cancer	Number	OMB Urban	RUCA Urban	RUCA Rural	State Incidence	Urban Incidence	Rural Incidence	National Incidence Rate	Nebraska SEER	SEER Inc Range by County
2005Peds	103	58	56	47	18.5501	18.64727	18.43564			
2004Peds	108	69	67	41	19.45059	22.31013	16.08215			
2003Peds	96	61	60	36	17.28942	19.97922	14.12091			
2002Peds	95	54	52	43	17.10932	17.31533	16.86665			
2001Peds	94	50	49	45	16.92922	16.31636	17.65114			
2000Peds	73	41	42	31	13.14716	13.98546	12.15968			
1999Peds	88	50	50	38	15.84863	16.64935	14.90541			
1998Peds	96	54	54	42	17.28942	17.9813	16.4744			
1997Peds	81	42	40	41	14.58794	13.31948	16.08215			
1996Peds	90	49	47	43	16.20883	15.65039	16.86665			
TotalPeds	924	528	517	407	16.64106	17.21543	15.96448	NA	NA	NA

Next Stage: Impacts of Socio-Economic Status

- For the purposes of comparison of cancer incidence to the potential characteristics of the patient as indicated by their home location, a socio-economic status (SES) index was developed.
  - This was done so as to approximate the most likely characteristics of the individual based upon the known social and economic characteristics of the local region as reported by the Bureau of the Census.
  - Block groups form the basis of the all analysis. Those block groups are then aggregated into the two or three Metro / Rural taxonomies used for the respective studies.

Nebraska Block Group
Socio-Economic Status Index



References

- Meza JL. Empirical Bayes estimation smoothing of relative risks in disease mapping. *Journal of Statistical Planning and Inference* 2003;112: 43-62.
- Besag JE, York J, Mollie A. Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* 1991; 43:1-59.
- Li J, Thompson D, Miller J, Pollack LA and Steward SL. Cancer incidence among children and adolescents in the United States, 2001-2003. Pediatrics -e2008; 121; e1470-e1477.
- U.S. Cancer Statistics Working Group (2004) United States Cancer Statistics: 2004 Incidence and Mortality. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer
- U.S. Department of Health and Human Services. Healthy People 2010: Understanding and Improving Health. Washington, DC: U.S. Department of Health and Human Services, 2000.
- Sam Harper and John Lynch (2005) Methods for Measuring Cancer Disparities: Using Data Relevant to Healthy People 2010 Cancer-Related Objectives http://seer.cancer.gov/publications/disparities/measuring_disparities.pdf

Funding Organizations

Nebraska Center for Rural Health Research.

- www.unmc.edu/rural

University of Nebraska Medical Center

- www.unmc.edu

Nebraska Department of Health and Human Services

- www.dhss.ne.gov

Contact Information

- Michael Shambaugh-Miller, Ph.D.
  - 402-559-7858
  - mdmiller@unmc.edu
- UNMC College of Public Health – Dept. of Health Services Research and Administration
  - 402-559-5260
  - www.unmc.edu/nebraska