QUANTIFYING CLASS INEQUALITIES IN CANCER:
CONCEPTS, MEASURES, METHODS AND CONTROVERSIES

Nancy Krieger
Harvard School of Public Health

Plenary Session: “Social class: theory and measurement”
North American Association of Central Cancer Registries Annual Meeting
Toronto, Canada, June 11-13, 2002
WHY CLASS & HEALTH?

• Growing recognition of social class as a critical determinant of population health

• At issue: multiple pathways, across the lifecourse

• Important empirical challenge:
  - how best to study & quantify connections between class & health, including cancer
  - obstacle: eclectic or absent socioeconomic data

• Unscientific—and not neutral—to ignore links between class & health
PURPOSE OF TALK

- Conceptual & methodological:
  - which measures, at which level & time period, for what purpose?
  - guiding perspective: ecosocial theory

- Empirical:
  - example: US cancer registries, geocoding & area-based socioeconomic measures
WHAT IS SOCIAL CLASS?

- Social relationship:
  - interdependent economic relationships among people (not individual attribute)
  - asymmetrical, premised on property & labor
  - one dimension of social position (cf race/ethnicity, gender, sexuality, etc.)

- Component of socioeconomic position
  - aggregate concept combining resource-based and prestige-based measures
  - “SES”: conflates the two and arbitrarily privileges status over material resources
EMBODIING CLASS

• Multiple pathways & levels:
  - “standard of living” (household, community, nation)
  - occupational exposures (paid & non-paid work)
  - environmental exposures (home, work, school, neighborhood, etc.)
  - social exposures (class discrimination)

• Multiple time periods:
  - pre-conception & in utero
  - childhood & adolescence
  - adulthood

• No “one size fits all” exposure-outcome relationship (cf lung vs breast cancer)
MEASURING SEP

- Diverse domains:
  - income
  - poverty
  - wealth (assets, including home ownership)
  - occupation & employment
  - education
  - crowding (>1 person/room)

- Single measures or combined?
  - if combined: assumes underlying construct (vs different measures perform differently)
CHOOSING MEASURES

• Which measure to use:
  - Empirical question, not philosophical principle, to be driven by research question
    • monitoring, etiology, access to/quality of care?
  - At issue:
    • which measure(s), singly or combined?
    • which level(s), for which time period(s)
    • in relation to which other social inequalities in health?

• Poorly chosen measures: obscure more than they reveal (e.g., using only education to investigate contribution of socioeconomic inequality to racial/ethnic disparities)
GEOCODING AND MONITORING US SOCIOECONOMIC INEQUALITIES IN HEALTH: DOES CHOICE OF AREA-BASED MEASURE AND GEOGRAPHIC LEVEL MATTER?—the Public Health Disparities Geocoding Project

Nancy Krieger, Pamela D. Waterman, Jarvis T. Chen, SV Subramanian, Mah-Jabeen Soobader, Rosa Carson

Funded by: NIH/NICHD (1 R01 HD36865-01)
With acknowledgments to:

**Massachusetts Department of Public Health:**

- Dr. Daniel Friedman (Assistant Commissioner, Bureau of Health Statistics, Research and Evaluation)
- Alice Mroszczyk, Program Coordinator for 24A/B/111B Review Committee;
- **Cancer Registry:** Dr. Susan Gershman, Director; Mary Mroszczyk, Geocoding/Special Projects Coordinator; Ann R. MacMillan, Data Analyst;
- **Registry of Vital Records and Statistics:** Elaine Trudeau, Registrar of Vital Records; Charlene Zion, Public Information Office

**Rhode Island Department of Health:**

- **Vital Statistics:** Roberta Chevoya, State Registrar of Vital Records;
- **Division of Disease Prevention and Control:** Dr. John Fulton, Associate Director
**RATIONALE FOR STUDY**

- **Paucity** of socioeconomic data in US public health surveillance systems—limits capacity to monitor social inequalities in health

- **Possible solution**: geocoding & area-based socioeconomic measures

- **Problem**: no consensus on **which** area-based socioeconomic measures, at **which** level of geography (census block group, census tract, ZIP Code)

- **Our study**: empirical investigation to address this problem, across multiple health outcomes (mortality, cancer incidence, low birthweight, childhood lead poisoning, STDs, TB, non-fatal weapons-related injury)
CRITERIA FOR ABSMs

• External validity
  – do the measures find gradients in the direction reported in the literature?

• Robustness
  – do the measures detect expected gradients across a wide range of health outcomes?

• Completeness
  – is the measure relatively unaffected by missing data?

• User-friendliness
  – how easy is the measure to understand and explain?
## STUDY POPULATION: people

<table>
<thead>
<tr>
<th></th>
<th>Massachusetts</th>
<th>Rhode Island</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990 population</td>
<td>6,016,425</td>
<td>1,003,464</td>
</tr>
</tbody>
</table>

**Mortality data***

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

**Cancer data**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(primary invasive)</td>
<td>140,610</td>
<td></td>
</tr>
<tr>
<td>1988–1992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989–1992</td>
<td></td>
<td>19,808</td>
</tr>
</tbody>
</table>

*all-cause, plus analyses of top 5 causes by race/ethnicity: heart disease, malignant neoplasm, cerebrovascular disease, pneumonia and influenza, chronic obstructive pulmonary disease, unintentional injury, diabetes, HIV, and homicide and legal intervention.

*all-cause, plus analyses of breast, cervix, colon, lung, prostate
## STUDY POPULATION: areas

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>population size (SD)</td>
</tr>
<tr>
<td><strong>MA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block group</td>
<td>5,603</td>
<td>1,085.4 (665.2)</td>
</tr>
<tr>
<td>Census tract</td>
<td>1,338</td>
<td>4,571.8 (2,080.0)</td>
</tr>
<tr>
<td>ZIP Code</td>
<td>474</td>
<td>12,719.7 (12,244.1)</td>
</tr>
<tr>
<td><strong>RI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block group</td>
<td>897</td>
<td>1,137.7 (670.8)</td>
</tr>
<tr>
<td>Census tract</td>
<td>235</td>
<td>4,325.3 (1,810.9)</td>
</tr>
<tr>
<td>ZIP Code</td>
<td>70</td>
<td>14,335.2 (13,234.8)</td>
</tr>
</tbody>
</table>
**ABSMs**

- **Occupational class**
  - working class; unemployment
- **Income**
  - median household; low; high, Gini coefficient
- **Poverty**
- **Wealth** (expensive homes)
- **Education**
  - low (< high school); high (>= 4 yrs college)
- **Crowding**
- **Composite**
  - Townsend index; Carstairs index; Index of Local Economic Resources
  - SEP 1; SEP 2; factor 1; factor 2; SEP index
  
  note: missingness typically <1% (but wealth: 4%)
# Geocoded Records (MA)

% of records geocoded to:

<table>
<thead>
<tr>
<th></th>
<th>Mortality (n = 156,366)</th>
<th>Cancer Incidence (n = 166,370)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block group</td>
<td>93.8</td>
<td>92.2</td>
</tr>
<tr>
<td>Census tract</td>
<td>99.8</td>
<td>100.0</td>
</tr>
<tr>
<td>ZIP Code*</td>
<td>99.9</td>
<td>100.0</td>
</tr>
<tr>
<td>not geocoded</td>
<td>0.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

* of the total MA & RI records geocoded to the ZC level, 6.3% could not be linked to 1990 census ZC data (non-residential ZC or ZC created or changed after 1990 census)
<table>
<thead>
<tr>
<th></th>
<th>Rate: least resources</th>
<th>Rate: most resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSM</td>
<td>BG 13.1 CT 14.4 ZC 11.7</td>
<td>BG 6.3 CT 6.9 ZC 4.2</td>
</tr>
<tr>
<td>Working class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median HH income</td>
<td>BG 13.8 CT 14.2 ZC 11.8</td>
<td>BG 6.4 CT 7.1 ZC 5.7</td>
</tr>
<tr>
<td>Poverty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>BG 11.0 CT 12.4 ZC 9.7</td>
<td>BG 7.7 CT 8.8 ZC 8.5</td>
</tr>
<tr>
<td>Wealth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crowding</td>
<td>BG 17.0 CT 10.0 ZC 21.8</td>
<td>BG 8.3 CT 9.1 ZC 8.2</td>
</tr>
<tr>
<td>Low education</td>
<td>BG 15.0 CT 14.0 ZC 11.9</td>
<td>BG 7.0 CT 7.8 ZC 6.4</td>
</tr>
<tr>
<td>Townsend</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ILER</td>
<td>BG 14.8 CT 15.5 ZC 12.9</td>
<td>BG 6.5 CT 7.0 ZC 5.5</td>
</tr>
<tr>
<td>SEP1</td>
<td>BG 17.6 CT 15.1 ZC 13.8</td>
<td>BG 6.3 CT 6.7 ZC 4.3</td>
</tr>
<tr>
<td>SEP index</td>
<td>BG 14.5 CT 15.2 ZC 12.3</td>
<td>BG 6.2 CT 6.9 ZC 5.4</td>
</tr>
<tr>
<td>Median value</td>
<td>BG 14.5 CT 14.4 ZC 11.9</td>
<td>BG 6.4 CT 7.1 ZC 5.7</td>
</tr>
</tbody>
</table>
# CERVIX: RII (MA)

**Relative index of inequality (95% CI)**

<table>
<thead>
<tr>
<th>Measure</th>
<th>ABSM</th>
<th>BG</th>
<th>CT</th>
<th>ZC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working class</td>
<td>2.6</td>
<td>2.4</td>
<td>2.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Median HH income</td>
<td>2.4</td>
<td>2.3</td>
<td>2.4</td>
<td>2.3</td>
</tr>
<tr>
<td>Poverty</td>
<td>2.7</td>
<td>2.3</td>
<td>2.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Gini</td>
<td>1.5</td>
<td>1.5</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>Wealth</td>
<td>2.3</td>
<td>2.4</td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td>Crowding</td>
<td>3.5</td>
<td>3.1</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Low education</td>
<td>2.5</td>
<td>2.3</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Townsend</td>
<td>2.8</td>
<td>2.4</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>ILER</td>
<td>2.7</td>
<td>2.5</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>SEP1</td>
<td>3.0</td>
<td>2.6</td>
<td>2.9</td>
<td>2.9</td>
</tr>
<tr>
<td>SEP index</td>
<td>3.0</td>
<td>2.6</td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td>Median value</td>
<td>2.6</td>
<td>2.4</td>
<td>2.7</td>
<td></td>
</tr>
</tbody>
</table>
Scaled RII Plot (BG)
KEY FINDINGS

• Choice of both area-based socioeconomic measure (ABSM) and geographic level matters

• Level of geography:
  − similarity of block group & census tract expected
  − worth effort compared to use of potentially more misleading ZIP Code data (and no ZIP Codes in 2000 census, only ZCTA)

• Choice of area-based socioeconomic measures:
  − Overall: robust, but some evidence of different measures more relevant for different pathways
  − Measures of economic deprivation: most sensitive, across all outcomes
CAVEATS & COMPARISONS

• Sources of error & bias
  – Geocoding & underregistration of cases: if associated with poverty, then a conservative bias
  – ABSM: very small % missing data
  – Temporal: simultaneity ok for monitoring (burden of disease; not same as etiologic research)
  – Spatial correlation: precision (vs adjacency)
  – Ecologic fallacy: not relevant

• Comparison to prior studies
  – None directly comparable
  – US research: 6 studies, focus on individual vs ABSM methods (BG & CT more consistent than ZC)
  – UK research: different area-based measures of deprivation, but at same level of geography
  – No studies: stratified by race/ethnicity & gender
CHOOSING AN ABSM: CRITERIA

• Based on:
  - our *a priori* criteria (*external validity; robustness; completeness; user-friendliness*), and
  - desirable attributes of an indicator (*Rossi & Gilmartin, 1980*)
    • conceptually-based
    • constructed from valid, reliable, and accessible data using appropriate statistical techniques
    • comparable over time and across population groups; and
    • readily understandable, with normative value relevant to timely policy making
TENTATIVE CONCLUSION

- Efforts to monitor US socioeconomic inequalities in health using area-based socioeconomic measures will be best served
  - by those tract or block group measures that are
    - most attuned to capturing economic deprivation
    - meaningful across regions and over time, and
    - easily understood
  - hence based on readily interpretable variables with a priori categorical cut-points

- Likely candidate:
  - US census tract data on % below poverty
REFERENCES:
Geocoding & area–based socioeconomic measures


REFERENCES:
Geocoding, class & cancer

REFERENCES:

social class & health—concepts & measures


REFERENCES:
ecosocial & epidemiologic theory

Extra slides
ALL-CAUSE MORTALITY (MA)

BG: % below poverty

BG: Townsend index

BG: SEP1 (poverty, occupation, and wealth)

QUESTION: which best conveys information?
## MEDIAN RII (MA)

### Median relative index of inequality

<table>
<thead>
<tr>
<th>Outcome</th>
<th>BG</th>
<th>CT</th>
<th>ZC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cancer incidence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites</td>
<td>0.9</td>
<td>0.9</td>
<td>1.2</td>
</tr>
<tr>
<td>Lung</td>
<td>1.4</td>
<td>1.3</td>
<td>1.4</td>
</tr>
<tr>
<td>Breast</td>
<td>0.8</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>Cervix</td>
<td>2.6</td>
<td>2.4</td>
<td>2.7</td>
</tr>
<tr>
<td>Prostate</td>
<td>0.6</td>
<td>0.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Colon</td>
<td>0.9</td>
<td>0.9</td>
<td>1.2</td>
</tr>
</tbody>
</table>

### Cancer mortality

<table>
<thead>
<tr>
<th>Outcome</th>
<th>BG</th>
<th>CT</th>
<th>ZC</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sites</td>
<td>1.2</td>
<td>1.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>