LECTURE # 4
DATA: USES, ABUSES & ANALYSIS

MARCIA CASTRO
Department of Global Health and Population
Harvard School of Public Health
mcastro@hsph.harvard.edu

Objectives

• Describe types of data available for health-related research
• Introduce basic demographic concepts
• Describe spatial analytical approaches
  – GIS, Remote Sensing, and Spatial Analysis
• Policy: Translating findings of multidisciplinary research

A systemic view of diseases

• Local characteristics & determinants & perceptions, result in unique patterns of disease transmission, require custom-made interventions, and impact population in different ways

One size does NOT fit all
Data Sources

- Three types of data:
  - Events
    - Cases & deaths
  - Exposure ("denominators")
    - Population
  - Correlates of transmission mechanisms
    - Multidisciplinary: social, economic, demographic, behavior & individual perception/knowledge, ecological, political, biological

Data Sources

- What are we looking for?
  - Research question
  - Time / Space / Population groups
  - Scale
  - Primary vs secondary

you can't always get what you want

Data Sources: Events

- Administrative data
  - Outpatient / Hospitalization
  - Reported cases
    - Active / Passive
  - Immunization
  - Deaths
  - Control activities

- Survey-based data
  - Sample frame
**Data Sources: Events**

- **Main sources**
  - Health care – in-patient, immunization, family care, and nutrition
  - Morbidity/Epidemiology – diseases that require compulsory notification, out-patient by cause
  - Data on HIV/AIDS - [www.aids.gov.br](http://www.aids.gov.br)
  - Mortality – data from death registration, ICD10

**Data Sources: Denominators**

- **Population counts**
  - Census
    - Every 10 years
  - Civil registration system
    - Coverage
  - Household surveys
    - Sample
    - Periodicity & Coverage
  - Projections
    - Uncertainty

**Data Sources: Denominators**

- **Main sources**
  - Brazilian Census Bureau (IBGE):
    [http://www.ibge.gov.br](http://www.ibge.gov.br) – census, household surveys, population projections
  - National Survey on Demography and Health of children and women (PNDS-2006)
Data Sources: Multidisciplinary

- Census & household surveys
- Agriculture
- Environment
  - Land cover, land change, land use
  - Climate
  - Soil, hydrology
- Infrastructure (e.g., roads, health facilities)
- Urban planning & expansion
- Social programs (“cash transfers”)

Data Sources: Multidisciplinary

- Main sources
  - Brazilian Census Bureau (IBGE): [http://www.ibge.gov.br](http://www.ibge.gov.br) – agrarian census, price indices, sanitation, urbanization, etc
  - National Institute of Spatial Research (INPE): [http://www.inpe.br/](http://www.inpe.br/) - deforestation, weather and climate, forest fires, satellite images
  - Data on cash transfers (Bolsa Familia) and other social programs: [http://aplicacoes.mds.gov.br/sagi/FerramentasSAGI_menu/internet.php](http://aplicacoes.mds.gov.br/sagi/FerramentasSAGI_menu/internet.php)

A word of caution

- Combining data …
  - … from different sources
  - … from different years
  - … in different scales
- Ecological X Individual analysis
- Generalizability
- Association vs Causality
Demographic Concepts

- Population (P) numbers can only change by
  - Births (B)  *Fertility*
  - Deaths (D)  *Mortality*
  - Movements in (I) / out (E)  *Migration*

- Measuring each component
  - Direct methods
  - Indirect methods:

Demographic Concepts  
*Mortality*

- Key indicators:
  - Infant & Child mortality
  - Life expectancy at birth

Demographic Concepts  
*Fertility*

- Key indicator: Total Fertility Rate

Demographic Census, 1991, 2000 and 2010
Demographic Concepts

Migration

- Migratory flow
  - Direction
  - Intensity
  - Characteristics

Spatial Methods

Space is special

- Visualization
  - Generate new variables
  - Sampling

- Exploration
  - Clustering
  - Early warning / surveillance systems

- Modeling
  - Explanatory / predictive models
  - Targeted interventions

Visualization: Visceral Leishmaniasis, Pernambuco

Visualization: Age/sex-standardized municipal tuberculosis notification rates per 100,000 in Brazil 2002–2009

Visualization: Vulnerability to dengue, São Paulo, July 2011-June 2012
Patterns?

Virgin Mary in a pancake
Mermaid in clouds

Exploration: clustering

Machadinho Settlement Project, Rondônia State, Brazilian Amazon

Low rates
Moderate rates
High rates

Significant clusters of low malaria rates
No significant clusters
Significant clusters of high malaria rates

Castro et al, 2006
Modeling: predictive model

Modeling: spatial diffusion

Translational research
If the analysis is multidisciplinary, so are the results and the translation

- Intersectoral collaboration
  - Public health, urban planning, transportation, agriculture, mining, education
- Synergy between different disease control programs
  - Similar vectors
  - Similar determinants/causes
References


• Kenneth Hill, Alan D Lopez, Kenji Shibuya, Prabhat Jha, on behalf of the Monitoring of Vital Events (MoVE) writing group. 2007. Interim measures for meeting needs for health sector data: births, deaths, and causes of death. The Lancet 370(9600): 1726-1735.
Questions for Discussion

• What are the main data limitations for health-related analysis?
  – What would be needed to fill in the gap(s)?
• How could available data be used for evidence-based planning of interventions?
• Is there a balance between local x global analysis?
  – How can each approach inform policy makers?
  – Which is most appropriate to evaluate the progress toward achieving the MDG?
• Other questions