Exposure Science: A Tool for Assessing Public Health Impact

Dan Wartenberg, PhD
Environmental and Occupational Health Sciences Institute
UMDNJ—RW Johnson Medical School, Piscataway, NJ 08854

dew@ehsi.rutgers.edu
The Take-Home Message

• Aggregate data are used routinely and successfully in environmental epidemiology
  – Aggregate data are individual or combined values used to represent a region, time period, population or group of exposures

• Ecologic analyses are valid assessments of exposure—disease relationships that have helped us screen and prioritize observations and demonstrate the utility of public health intervention
  – Ecologic analyses are those in which we do not have the complete joint distribution of exposure and disease

• Aggregate data and ecologic analyses are one part of a continuum, from populations to individuals to systems within individuals, all of which contribute important information to our understanding of environment and health
EPA’s “Environmental Public Health Continuum”

Modified from EPA RFA
Our Charge

• In the last 10 years…..
  why hasn’t “Exposure Science” research
  …..progressed as far as it could?

  – Consider
  • The State of the Science
  • Establishing better linkages across the continuum
  • Identifying barriers along the continuum

My Charge

Public Health Monitoring and Tracking: Making Sense of Exposure

Is this session title appropriate:
  “Human Contact (Dose) / Internal Dose (Biomarkers)”???
Relationship to “Disease Tracking”

• Pew Recommendation
  – “Tracking” is synonymous with CDC’s concept of public health surveillance (Health Gap, Sept. 2000)

• CDC’s Environmental Public Health Tracking Program
  – Mission: To provide information from a nationwide network of integrated environmental monitoring and public health data systems so that all sectors may take action to prevent and control environmentally related health effects (Qualters, 2004)

• An Epidemiologist’s Definition of Surveillance
  – A public health disaster is an event so severe that even an epidemiologist can detect it!
    -- David Ozonoff, BUSPH
My Talk

• **The context**
  – Identifying existing data sets of ambient exposure levels
  – Undertaking analyses of aggregate data (concepts/methods)

• **The exposure continuum**
  Emission → Environment → Biomarker
  Regulation → Public Health → Medicine

• **The application**
  – Large scale, long term, population-based, studies and interventions
    • Public health impact investigations can be initiated based on **disease burden** or **exposure burden**.
Two Approaches

Exposure----------------- Disease
Example: Air Pollution------- Lung Disease
          Air Pollution------- Heart Disease

Disease------------------- Exposure
Example: Lung Disease------- Air Pollution
          Heart Disease------- ???

Would heart disease have been identified as a consequence of exposure to air pollution?
Some Types of Aggregate Data

• Generic
  – “Indicators”
    • Summary measures that capture multiple aspects of a situation of concern (e.g., %low birthweight, extent of wetlands, concentration of criteria pollutants, beach closings)

• Exposure
  – Substantive aggregation
    • Measure of effective exposure (e.g., dioxin congeners)
  – Regional measurement
    • May be due to monitoring or reporting system
      – Regional exposure (e.g., pesticide use, taxes)
      – Temporal exposure (e.g., average annual, 24-hour average)
      – Sampling limitations (e.g., ambient air monitors)

• Outcome
  – May be aggregated to protect privacy (e.g., registries, surveys)
Some Major Environmental Health Concerns

Where have we been…

…and where do we want to go

Exposures often characterized by aggregate measures

- Air quality
- Drinking water quality
- Lead
- Ionizing radiation
- Magnetic fields
- Climate change

Exposures often characterized for individuals

- Radon
- Cell phones
- Pesticides
Analysis of Aggregate Data

• Ecologic studies “don’t get no respect”
• But, it ain’t so:
  - Despite limitations, not without virtue
  - New technologies can refine exposure estimates using multiple data sources—e.g., GIS, satellites, land use, measurements
  - Improved yet simple analytic tools enable VALID assessment of association despite effect size bias
  - Typically population-based—increases Public Health Relevance
  - Provides broad look at less well understood issues
    - Better space-time coverage, larger sample size
  - Especially useful for assessing interventions
• One of many tools, all of which have their place in the public health study tool kit
Why have Ecologic Studies been given a bad name?

- Misapplication
  - Rarely preferable when individual data are available
- Misuse
  - Failure to consider bias, confounding, effect modification (similar to individual-based studies)
- Misinterpretation
  - Lack of familiarity with methodology may lead users to over interpret or extrapolate results
- Lack of joint distribution of exposure and outcome when both are aggregate

- But, usually blamed on aggregate Data
Goal: Exposure → Etiology

Consider Cancer Incidence

• **Ideal Case**
  – Know the amount (#molecules) of relevant toxic that enters a susceptible cell and causes change in DNA leading to disease over a lifetime

• **Realistic Case**
  – Have crude estimate of ambient or self-reported exposure to toxic, for a limited amount of time, through limited exposure routes, and some measure of disease occurrence
Missing Information

• Space-Time Trajectory
  – Full residential/work/travel history
  – Complete list of exposures for each location at each time over lifetime

• Risk Factor History
  – Diet
  – Behavior (drinking, pharmaceuticals)
  – Occupation/Hobbies
Strengths of Aggregate Analysis

• Enables analysis of large populations
  – Not easily collectable
  – Facilitates study of relatively small risks
  – Can assess public health impact of an intervention
  – Can be conducted easily and inexpensively with routinely collected databases (surveillance)

• *Statistical inference is valid in spite of biased estimates* (*manuscript under revision*)
  – Useful for hypothesis generation and prioritization
  – Aggregate sampling variance biases results towards null
Where Have We Come From

- **Air pollution epidemiology**
  - Severe episodes (Meuse Valley, Donora PA, London Fog)
  - Time series studies (daily mortality in Philadelphia)
  - Regional comparisons (e.g., Six Cities, ACS)

- **1975 NCI Cancer Mortality Atlas**
  - Found new and confirmed known etiologies
  - Validation slow, partially successful but fruitful
  - Even occupational risks identified by geographic aggregation

- **Migrant study (e.g., diet, cancer mortality)**
  - Do transplanted populations acquire disease rates of local populations?
    - Convenient sample, integrated exposure
    - “Descriptive” or “Analytic”
Air Pollution

• 1st Generation Studies
  – Obvious and extreme exposure
  – Severe and acute outcomes
  – *Ecologic Design*

• 2nd Generation Studies
  – Many comparison of health outcomes across exposure gradients (6 cities; ACS)
  – Some *Ecologic Designs*
  – *Led, in part, to Clean Air Act*

• 3rd Generation Studies
  – More fine scaled exposures and outcomes
  – Some times series work uses *Ecologic Designs*
Map-Based Correlational Studies

- Various historical efforts
- New impetus triggered by NCI Atlas (1970s)
  - Compared mortality maps to possible exposures
  - Then validated with traditional epidemiology
    - Bladder cancer and chemical manufacturing
    - Nasal adenocarcinoma and furniture manufacturing
    - Lung cancer and shipyards
    - **Oral cancers among women and snuff use**
- Despite the ‘Bad Press’ these can be useful
- Must be careful of limitations of ‘ecologic analysis’
My Main Message

• Each part of the continuum has value
• Each part of the continuum has limitations
• Through integration of all, we can best
  – Identify and prioritize etiologic research
  – Develop cures and interventions
  – Implement effective interventions
Challenges

• Study Design
  – Population vs. Individual base
    • Public health vs. Medicine
    • Practicality: sample size, cost, duration
  – Initiated from exposure vs. disease concern
  – Etiologic, Screening or Intervention Goal
    (i.e., medicine vs. public health)

• Analytic Methods
  – Individual vs. Aggregate data
    • Bias, validity and generalizability
  – When does our exposure data outstrip our health outcome data?
The Take-Home Message

• Aggregate data are used routinely and successfully in environmental epidemiology
  – *Aggregate data* are individual or combined values used to represent a region, time period, population or group of exposures

• Ecologic analyses are valid assessments of exposure—disease relationships that have helped us screen and prioritize observations and demonstrate the utility of public health intervention
  – *Ecologic analyses* are those in which we do not have the complete joint distribution of exposure and disease

• Aggregate data and ecologic analyses are one part of a continuum, from populations to individuals to systems within individuals, all of which contribute important information to our understanding of environment and health
Exposure Meets Epidemiology’s Limits

• Often, epidemiologic data are incomplete
  • Behavior, diet, residence history, work history, previous exposures
  • Genetics
    – Outcome data may be crude (Y/N)
• But, exposure models can be extremely precise
  – may not be accurate
  – measures may not capture full exposure experience
• How should we handle these situations?
  – Reduce precision to reflect accuracy?
  – Try to match accuracy and specificity across health/exposure realms?
  – Target funding/research at how best to address impact of the issue under study rather than doing so for each component independently
The Environment Controversy
“arguments against”

• Environment does not cause much disease
  – Spatial patterns are due to other factors
• We do not have adequate spatial data
  – Lifetime residence history
  – Other exposures (e.g., occupation, behavior)
• Methods are inadequate
  – Ecologic fallacy (aggregation bias)
The Environment Controversy “arguments for”

- Many environmental exposures are spatially cohered
- Location matches on many unmeasurable factors (e.g., ethnicity, behaviors, SES)
- Location can help target specific populations
  - Extremes of occurrence or exposure gradients
- Spatial patterns can generate hypotheses for further testing (acknowledging methodologic limits)
Existing Data Sets of Ambient Exposure Levels

• Goal of data collection is to characterize the environment
  – Not linked to
    • Specific individuals
    • Particular events
  – Typically used as regional measures

• Examples
  – Air quality monitors (sparse)
  – Groundwater assessments
  – Drinking water samples (few per system)