Environmental Epidemiology

Faith G Davis, PhD, FACE
Professor and Vice-Dean
What is Environmental Epidemiology?

• Epidemiology is the study of the distribution and determinants of disease in the population

• Environmental epidemiology studies the effects of environmental exposures on health and disease in the population

• Environmental health? all that surround us
Reasons to embark on studies

• Looking for a cause of disease
  – Knowledge of disease mechanisms may help formulate which environmental exposure to examine.

• Concern for environmental factor that may lead to disease.
  – Knowledge about toxicity or harmfulness of the environmental factor may help formulate hypothesis.
Studies: also a response to

- Chemical incident
- Natural disaster
- Ecologic disaster

- In general, risk estimates
  - inform cost-effectiveness analysis
  - Inform policy decision
Cholera death rates in London

Lessons learned (1853)
- tenants did not know their water supply source (tested water salt content to determine)
- used geography to illustrate relationships
Other Examples

• London Smog (1953)
  – Acute – mortality increased almost immediately
  – Took weeks to return to Normal death rates
• Arsenic in well water in Bangledesh (1970s)
  – Skin lesions became apparent in about 10 years
  – Cancer rates have now started to increase
• Bhopal chemical (MIC) spill (1984)
  – Immediate deaths from choking
  – Longterm chronic effects in survivors
How Do Hypothesis Develop?

• Astute Clinicians
  – 8 cases of adenocarcinoma of the vagina

• Observing trends
  – increase in lung cancer in the 1930s
  – increase in endometrial cancer in the 1970s
  – decline in stomach cancer

• Previous studies
  – epidemiologic or biologic
Credible Hypothesis (educated guess)

Incorporate all available knowledge in framework of causal criteria continuously modify with new knowledge while avoiding thoughtless/needless repetition
Achilles Heel of Env Epi

• Exposure Assessment
• Strategy must match knowledge of agent, its interaction with humans, health effect, study design and budget.
• Goal – accurate, precise, biologically relevant, for the critical exposure period, show a range of exposures
Exposure and Dose

- **Exposure** – contact of a substance in an environmental medium (water, air, soil) and the surface of the human body (skin, respiratory tract).
- **Dose** – the amount of the hazard that enters the body.
- **Target organ dose** – the amount of the agent that reaches the susceptible organ.
Exposure Considerations

- **Agents:** Chemical, Biological, Physical
- **Media (vectors):** water, air, soil, food
- **Routes:** inhalation, ingestion, absorption
- **Parameters:** duration, concentration, frequency
- **Measures:** direct, indirect or surrogate
- **Consider genetic variation (susceptibility)**
- **Confounders**
Study Design Considerations

- Acute versus chronic effects – short term or cumulative exposure, respectively
- Individual vs group measure – access, cost and precision tradeoffs
- Classification - dichotomous, ordinal
- Expert, self assessment vs measurement
- Modeling – deterministic, stochastic, GIS
- Validation
Exposure measure affects Study Design

• Increasing the population size could allow for cruder exposure estimates while
• Smaller population sizes require more refined exposure estimates.
• Balance sample size and cost with access to subject, available tools and data.
• If using routine health outcome data, modeling based on residence only
Heirarchy of Exposure Data

• Quantified personal measurement (Best)
• Quantified area measurement – near
• Quantified surrogates (estimates)
• Distance and duration of exposure
• Distance or duration of residence
• Residence or employment in area in reasonable proximity to site of exposure
• Residence or employment in a defined area (ie county) of the site of exposure. (Worst)
Counties with Potential Elevated Radon

EPA Map of Radon Zones

The purpose of this map is to assist National, State, and local organizations to target their resources and to implement radon-resistant building codes. This map is not intended to be used to determine if a home in a given zone should be tested for radon. Homes with elevated levels of radon have been found in all three zones. All homes should be tested regardless of geographic location.

IMPORTANT: Consult the EPA Map of Radon Zones document (EPA-402-R-93-071) before using this map. This document contains information on radon potential variations within counties. EPA also recommends that this map be supplemented with any available local data in order to further understand and predict the radon potential of a specific area.

Guam - Preliminary Zone designation
Distributions of Radon Measures and their Association with Gliomas

<table>
<thead>
<tr>
<th>Radon Measure</th>
<th>Cases (n=504)</th>
<th>Controls (n=804)</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>116 (23%)</td>
<td>50 (6%)</td>
<td>5.4</td>
<td>3.7, 7.9</td>
</tr>
<tr>
<td>Moderate</td>
<td>187 (37%)</td>
<td>266 (33%)</td>
<td>1.7</td>
<td>1.3, 2.2</td>
</tr>
<tr>
<td>Low</td>
<td>201 (40%)</td>
<td>488 (61%)</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

*using Multivariable Logistic Regression Models controlling for Age, Gender, and Race*
Exposure Measurement Error
Misclassification

- Differential error: bias in either direction
- Classical and Berkson Error: systematic and random components
- Classical – average of many replicate exposure of the same true exposure would equal the true exposure.
- Berkson – Proxy exposure is used for many subjects; the true exposure varies randomly around it with mean equal to it.
Consequences of Random Error

• For exposure on a numerical scale,
  • Classical errors bias RR towards 0 – error has to be relatively big to give serious bias (linear, logistic and log-linear models)
  • Berkson errors lead to no bias in linear regression and little to none in the others.

• Random Errors
  – Dichotomous - towards the null
  – Polytomous - downwards estimates of trends across ordered groups –between specific groups bias can be in either direction.
Additional Exposure Measure Considerations

- Errors in confounders and effect modifiers
- May be able to correct for errors
  - Validity studies
    - Sensitivity and specificity
    - Direct measure classical and Berkson errors.
  - Reliability studies
- Note: All types of error reduce study power
  - Can be measured if magnitude of error and exposure variability is known.
Classical Study Design Options

• Descriptive
• Ecological
• Cross-sectional
• Case-Control – rare disease, control selection, timing of exposure & other data
• Cohort - common disease, exp precedes
• Experimental – removal of exposures
Special Study Designs

• Time Series
  – Temporal variability - air pollution studies
• Panel Studies (time in common)
  – Effects on sensitive subjects
• Spatial Epi - Geographic variation
• Investigation of Disease Clusters – in response to perceived/real excess of cases
• Gene-environment interactions
Time Series

• Used to assess short-term changes in health following changes in exposure
• Routine E/D data aggregated over the same time (days) during a specified time period
• Confounding not as much of a problem - potential CFs not associated with temporal change in exposure
• Seasonal patterns at issue – best for time periods under 40 days.
Panel Studies

• Prospective studies that follow a small group of people intensively for a short time

• Individual repeated measures (every day):
  – exposure & outcome & confounders

• Maximize statistical power:
  – sensitive subjects (children, elderly)
  – may have limited generalizability
  – Evaluates the short term effects of time varying exposures (while cohort evaluated the effects of exposures on single incident event)

  – Analysis: aggregate or cohort with repeated data.
Spatial Epidemiology

• Disease mapping provides baseline data – change over time
  – Small area maps – latency and migration?
  – SIR/SMR – bayesian smoothing may address imprecision from small area estimation
• Geographic Correlation – group data
• Point source
  – Circular areas around source
  – Dispersion modelling (latency?)
Acanthamoeba keratitis (AK) case distribution overlying 2003 census tract population density, plotted by quartile (based on all AK cases diagnosed at the University of Illinois at Chicago from June 1, 2003, to May 31, 2007, in patients residing in the 5-county area and 2003 census population data). For the entire area, 2 or 3 cases are expected per year. Observed cases from June 1, 2003, to May 31, 2005, are distributed to the west, south, and southwest and farther from Lake Michigan compared with the expected population distribution. Comparison between periods demonstrates that during the period from June 1, 2005, to May 31, 2007, cases continued in the far west, south, and southwest but also occurred closer to Lake Michigan and the city center.
Disease Cluster Investigation

• Disease clusters: aggregations of similar or related diseases in groups
  – Individuals – seeking explanations from health authorities
  – Provide new clues to unknown etiology or even a new disease.

• Response, monitoring, research
  – Short term vs ongoing assessment
  – Hypothesis generation vs testing

• Managing public concern

• Debate – as most have not resulted in new information they are not a good use of resources
Gene Environment Interaction

• Paradigm – the vast majority of diseases arise because of gene-environment interactions
  – Monogenic – mendelian patterns
  – Complex genetic susceptibility – interaction between environmental risk factors and genes changes the risk of disease.
Fig. 1. Scheme of biomarkers of exposure, effect and susceptibility in environmental carcinogenesis.

Kyrtopoulos 2006

Exposure → Species or Individual Susceptibility → Clinical Disease

Biomarkers of Exposure:
- Internal dose
  - Chemicals / Metabolites / Genotoxic Activity
- Biologically Effective Dose
  - Protein Adducts / DNA Adducts

Biomarkers of Early Effect:
- Early Biological Effects
  - Gene Mutations / Microsatellite Instability / Cytogenetic Changes / Modified Gene Expression
- Altered Structure / Function
  - Mutation Spectra in Tumours or Pre-Cancerous Cells

Metabolism → DNA Repair and Other Responses to DNA Damage → Immune Defense

Susceptibility: Genetic / Acquired Variation
Exposure Precision – increase sensitivity & specificity to improve validity of measure

- Ecologic measures (group)
- Individual self-report (indirect)
- External dose (personal monitors)
- Internal dose (concentrations in tissue)
- Biologically effective dose (DNA adducts, hemoglobin adducts)
- Early biological response (point mutations, DNA repair genes)
Genotype is stable, measured accurately (sens, spec=90-100%), frequency of alleles is high

Environmental exposures are changing (life-course events), often measured inaccurately, frequency may be too low
In addition, genetic polymorphisms are investigated with high-throughput technologies that allow researchers to investigate hundreds of thousands of SNP at a time: with the usual p-values this originates a large number of false positives.

In environmental research false negatives are an important problem.
A self-fulfilling prophecy: are we underestimating the role of the environment in gene-environment interaction research?
(P Vineis Int J Epidemiol 2004)

According to estimates, the common genotyping method Taqman has 96% sensitivity and 98% specificity, thus allowing little error in classification. On the contrary, sensitivity in environmental exposure assessment is quite often lower than 70% and specificity even lower.

Relative Risks of 1.5 may be missed
Study Designs: difference aspects of E-D relationships

- **Time Series** – short term exposure disease relationships
- **Panel** – short term individual exposure disease relationships
- **Spatial** – is there a pattern by geography?
- **Cluster** – Is there an excess in this small region?
- **G-E** – Is there an association? Interaction?
<table>
<thead>
<tr>
<th></th>
<th><strong>Advantages</strong></th>
<th><strong>Limitations</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time Series</strong></td>
<td>Cheap and easy to apply</td>
<td>Data may not be optimal</td>
</tr>
<tr>
<td></td>
<td>Existing data - long time series</td>
<td>Individual variability cannot be studied</td>
</tr>
<tr>
<td></td>
<td>Ethics approval easy</td>
<td>Sensitive to modeling choice</td>
</tr>
<tr>
<td></td>
<td>Relatively free from confounding</td>
<td></td>
</tr>
<tr>
<td><strong>Panel Studies</strong></td>
<td>Individual measurements</td>
<td>Measures – affects compliance</td>
</tr>
<tr>
<td></td>
<td>Subclinical outcomes</td>
<td>Cost per subject high</td>
</tr>
<tr>
<td></td>
<td>Confounding addressed in analysis</td>
<td>Sample size decreases power</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complicated analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generalization??</td>
</tr>
<tr>
<td><strong>Spatial</strong></td>
<td>Explore unusual patterns and create new hypothesis</td>
<td>Data availability</td>
</tr>
<tr>
<td></td>
<td>Point source studies – may add to evidence of causality</td>
<td>Current exposure only?</td>
</tr>
<tr>
<td><strong>Clusters</strong></td>
<td>Exploratory/hypothesis testing</td>
<td>Inconclusive results, raise anxiety and consume resources</td>
</tr>
<tr>
<td><strong>G-E</strong></td>
<td>Identifying groups at high risk</td>
<td>Most assns weak</td>
</tr>
<tr>
<td></td>
<td>Increase plausibility of env assns</td>
<td>Assns heterogeneous</td>
</tr>
</tbody>
</table>
Paper Discussion

What is the study question?
What is the exposure measure?
  is the window of exposure appropriate?
Will error in exposure bias the results?
Do you agree with interpretation of results?
If you had been designing this study – what are the 2 things you would have liked to change?
Design Considerations

• Resources?
• Number of repeat exposure measures?
• Designing for maximum power
• Designing for Berkson rather than classical error
• Validity, reliability and 2-stage studies
Special Situations
Chemical Incidents/Natural Disasters

- Unexpected release of toxic material
- Change in physical environment caused by natural forces with marked adverse impact on human beings
- Study Design considerations:
  - Pre-event – anticipate and plan
  - During event – good record keeping
  - Post event (short-term) – exposure assessment, case identification, communication
  - Post event (long-term) – surveillance and long-term followup
Developing Countries

- Water and sanitation
- Indoor air pollution
- Outdoor air pollution
- Heavy metals
- Pesticides
- Persistent organic pollutants
- Malaria and parasitic diseases
- Susceptibility factors – poverty, malnutrition, crowding, poor living conditions
Research Settings - Rural

- Dispersion of population
- Lack of communication
- Cultural diversity
- Lack of relevant information: exposure and health data.
- Logistics of carrying out research

- Local census strategies
- Participatory discussions
- Alternatives for exposure modelling ie: satellite images
- Use easy validated existing field techniques
- Invest in transporting people and samples
Health Assessment

• Acute and chronic effects
• Specific and non-specific effects
• Individual variation and susceptibility
• Hyperreactivity and Hypersensitivity
• Measurement requires case definition
  – Existing records, questionnaires, physical exams, physiological measures, biological measures
Interpretation of Results

- Chance (probabilities)
- Error (Bias)
  - selection
  - classification
- Confounding
- True effect
“Negative” results in cohort studies (Hernberg 1981)

- Truly negative: large, sensitive, well-documented exposure data
- Small and/or insensitive studies are uninformative
- Design issues: crude measurement, wrong categories, subjects with too short (or too low) exposure, too short followup for latency, incomplete followup, wrong referent or statistics
- “Interpretation requires..knowledge of the subject and apprehension…that errors…tend to mask existing differences.”
Evaluation of results

• How large is the risk?
• Does risk increase as exposure increases?
• Does exposure proceed disease?
• Does removing exposure reduce disease?
• Are results consistent with scientific knowledge (animal and human)
To demonstrate Cause

- Reasonable number of persons
- exposed to a risk factor
- of some potency
- for some time
Ethical Considerations

• Respect for autonomy
• Beneficence
• Non-maleficence
• Distributive justice
• Precedent established from prior studies
• http://www.iseepi.org/about/ethics/html
Precautionary principle

• “in order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for post-postponing cost-effective measure to prevent environmental degradation” (UN 1992)
A framework for articulating precaution and risk assessment.

Andrew Stirling EMBO Rep. 2007;8:309-315
What is Environmental Impact on Cancer?

– 2% carcinogens in the environment
– 4% in the workplace (1981)
– “grossly underestimated” (Presidents Cancer Panel 2010)
– 7-19% toxic environmental exposure (IARC 2008)
– Multiple interacting factors involved so its impossible to assign percentages.
Halifax Project

• Call for reassessment of IARC programme on chemical safety “mode of action” framework
• “cumulative effects of individual (non-carcinogenic) chemicals acting on different pathways and a variety of related systems, organs, tissues and cells could plausibly conspire to produce carcinogenic synergies….Research focused on low-dose effects of chemical mixtures needs to be rigorously pursued (Goodson et al. Carcinogenesis. 2015)
“The fact that epidemiology deals with observations in humans is a source of strength. The obvious advantage is that this is the species of concern and the necessity to extrapolate inferences across species is avoided. A less widely recognized advantage is that people are numerous and they expose themselves to substances with abandon”
“Humans ..available in large numbers .. they house and feed themselves and keep themselves clean at no expense to the investigator. They choose a broad range of dosages to a variety of potentially toxic substances.”.. “Numbers have permitted the identification of innumerable hazards never suspected from laboratory experiments”
Acknowledgement

Environmental Epidemiology: Study Methods and Application
by Dean Baker and Mark Nieuwenhuijsen
Oxford Press 2008