Epidemiology: the Good, the Bad and the Truly Ugly

David M. Gute, Ph.D., M.P.H.
WSSS Seminar
October 12, 2007
Today’s Agenda

• “The Good, the Bad and the Truly Ugly”. Apologies to Clint.
• Selected examples from my career.
• Defining environmental epidemiology within a context of all epidemiology.
• Review of appropriate methods for conducting environmental epidemiology studies.
• Importance of surveillance (as a cross-cutting theme) and monitoring.
• Review of methods for ascertaining complex exposures and the controversies which can emerge (diet and water as examples)
• Questions and Discussion
Sorry Clint
Epidemiological/Public Health Triad

Host

Vector

Agent

Environment
My Wonderful Adventure: Selected Epi Applications (1)

• MPH thesis (1975-1976) examined the distribution of chronic obstructive lung disease in Maine. **Finding**: elevated rates in rural areas contrary to the usual elevations in urban areas.

• 1st public health position (1979-1983) in Rhode Island. **Finding**: resources skewed heavily towards curative medicine, geographic and specialty mal-distribution also present. $ for health care are sufficient—remedy = cap and re-allocate. **Action**: political defeat, grudgingly learned the wisdom of “incremental change”.
My Wonderful Adventure: Selected Epi Applications (2)

- Ph.D. Thesis (1978-1981)-With NIOSH funding evaluated the distribution of occupational mortality of RI decedents. **Finding:** Some elevations were anticipated (respiratory mortality and textile workers) while others were new and unanticipated (stomach cancer and female jewelry workers). **Action:** NIOSH, heartened by the surveillance potential of vital records, presses ahead with a national program.
My Wonderful Adventure: Selected Epi Applications (3)

• Assistant Commissioner in Massachusetts: (1983-1988) administered investigations which spanned the study of PCBs in New Bedford harbor and the cluster of childhood leukemia in Woburn.

• Life on the Hill (1988-present) Together with Richard Enander, Ph.D. characterize the autobody refinishing industry in RI. **Finding:** identify lead as a principal component in top selling body filler compounds. **Action:** lead removed from product line.
The Nature of Environmental Exposure

- Population less well defined: Why?
- Variable susceptibilities of population
- Database integrity and accuracy?
- Develop additional items through discussion
Acute Air Pollution Incidents

1948: Donora, PA at noon.
Smog: Not Just a Left Coast Problem

Credit: Neal Boenzi/NYT Pictures

1966: New York City buried under a sea of smog.
Emerging Infectious Diseases, Volume 9, Number 5, May 2003
Estimating the Incidence of Typhoid Fever and Other Febrile Illnesses in Developing Countries
John A. Crump,* Fouad G. Youssef,† Stephen P. Luby,* Momtaz O. Wasfy,† Josefa M. Rangel,* Maha Taalat,† Said A. Oun,‡ and Frank J. Mahoney*†
*Centers for Disease Control and Prevention, Atlanta, Georgia, USA; †U.S. Naval Medical Research Unit No. 3, Cairo, Egypt; and ‡Ministry of Health and Population, Cairo, Egypt
Methodological Features of Environmental Epidemiology

• Ecologic studies
  “Homogeneity of exposure within groups is in general desirable in ecologic studies, as is heterogeneity of exposure between groups, because these conditions make it easier to detect effects of exposure”.

• Clusters
  “The first problem is determining whether the cluster really represents a high occurrence of disease”.
A second "problem that arises is defining the boundaries of the area or time frame for the cluster to be used in calculating rates".

Rothman’s negative critique of cluster investigation- what are these attributes?

“Useful scientific data on environmental risks, on the other hand, is more likely to be found by determining good studies to investigate specific etiologic hypotheses, in populations in which no clusters has been noted.”
Methodological Features of Environmental Epidemiology (3)

- Errors in measuring exposure:
  “In many epidemiologic studies, it is very difficult or even impossible to measure the true exposure exactly. Hence the observed exposure is measured with error (for continuous variables) or misclassified (for categorical variables).

- The general presumption of nondifferential exposure measurement error or misclassification leading to bias towards the null may not hold.
Two Common Scenarios in High Income Nations

• Relatively high concentration of pollutants, plausible if not actual routes of exposure to human populations, chronic exposure. Examples: Superfund sites such as the two sites in Woburn, MA, Department of Energy sites and military bases in the US.

• Relatively low concentration of pollutants, less likelihood of exposure pathways to human populations, chronic exposure. Examples: Brownfield sites in the US.
Risk Assessment

- Risk assessment addresses:
  - Does the weight of the evidence indicate that exposure causes disease? *Hazard Identification*
  - If so, what is the quantitative dose-response relationship? *Dose-Response Assessment*
  - What does the relationship imply for regulation? *Estimates of the prevalence of exposure taken together with risk*

- Inputs include animal and human data.
Dose-Response Assessment Using Published Data

• Inputs:
  – Linear regression of SMRs (thus we need access to cohorts)
  – Cumulative exposure estimates (emphasis on retrospective exposure)

• Output:
  – Unit risk estimate (URE) \( \text{ug/m}^3 \) for inhalation and is defined as an estimate of the increased cancer risk from a lifetime exposure to a concentration of one unit exposure.
Epidemic Curves (1)

Epidemic Curve (non-propagating)

Cases per 10,000

Month

Jan Mar May Jul Sep Nov Jan
Point-Source Exposure

Number of Cases

Days

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

13579 1 1 1 3 1 5
Natural History of a Disease

• The progression of signs and symptoms of the disease under study.
• Often these stages offer opportunities for prevention and or detection.
• Knowledge of natural history also informs prognosis both in the individual and for the population.
Meta-Analysis: Another Important Component of Environmental Epidemiology

- Width of the confidence intervals are more evident.
- Application of homogeneity and heterogeneity demonstrated.
Basic Epidemiologic Study Designs

- Cluster or case reports
- Ecologic
- Cross-sectional
- Cohort Studies (prospective and retrospective)
- Case-Control Studies (Nested + Pop. Based)
- Longitudinal
- Random Control Trial (how would we mount these for asbestos exposure and stay out of jail?)
# Epidemiological Study Designs

<table>
<thead>
<tr>
<th>Unit of Study</th>
<th>Study Design</th>
<th>Also Called</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations</td>
<td>Observational studies</td>
<td>Descriptive and analytical studies</td>
</tr>
<tr>
<td>Individuals</td>
<td>Ecological</td>
<td>Correlational</td>
</tr>
<tr>
<td>Individuals</td>
<td>Cross-sectional</td>
<td>Prevalence</td>
</tr>
<tr>
<td>Individuals</td>
<td>Case-Control</td>
<td>Case-reference</td>
</tr>
<tr>
<td>Individuals</td>
<td>Cohort</td>
<td>Longitudinal, Follow-up</td>
</tr>
<tr>
<td>Patients</td>
<td>Experimental studies</td>
<td>Intervention studies</td>
</tr>
<tr>
<td></td>
<td>Randomised control trials</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Field trials, community</td>
<td></td>
</tr>
</tbody>
</table>
## Principal Sources of Human Health Data

<table>
<thead>
<tr>
<th>Mortality data</th>
<th>Complete, accessible, relatively inexpensive</th>
<th>Accuracy of diagnosis, latency problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morbidity data</td>
<td>Minimizes latency, relatively complete</td>
<td>Comparability, nonspecificity, expensive, accessibility</td>
</tr>
<tr>
<td>Hospital discharge data</td>
<td></td>
<td>Completeness, practicality, confidentiality</td>
</tr>
<tr>
<td>Reportable diseases</td>
<td>Timely, close to exposure</td>
<td></td>
</tr>
<tr>
<td>Incidence cancer registry (MA as example)</td>
<td>Improved quality of information, decreases latency, less costly</td>
<td>No outcome data</td>
</tr>
<tr>
<td>Follow Back cancer registry (CT as example)</td>
<td>Same as above except for cost</td>
<td>More expensive</td>
</tr>
</tbody>
</table>
Descriptive Epidemiology

• Concerned with description rather than explanation
• Looks for associations between cases based on person, place and time considerations:
  – Person - demographics, symptoms
  – Place - location, common exposures
  – Time - onset of illness, epidemic curve
### Person: Frequency Tables

**Distribution of age among all respondents**

<table>
<thead>
<tr>
<th>AGE</th>
<th>Freq</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>1</td>
<td>2.9%</td>
<td>2.9%</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>5.7%</td>
<td>8.6%</td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>11.4%</td>
<td>20.0%</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>8.6%</td>
<td>28.6%</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>8.6%</td>
<td>37.1%</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>5.7%</td>
<td>42.9%</td>
</tr>
<tr>
<td>27</td>
<td>3</td>
<td>8.6%</td>
<td>51.4%</td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>5.7%</td>
<td>57.1%</td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>2.9%</td>
<td>60.0%</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>2.9%</td>
<td>62.9%</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>2.9%</td>
<td>65.7%</td>
</tr>
<tr>
<td>34</td>
<td>3</td>
<td>8.6%</td>
<td>74.3%</td>
</tr>
<tr>
<td>35</td>
<td>5</td>
<td>14.3%</td>
<td>88.6%</td>
</tr>
<tr>
<td>36</td>
<td>1</td>
<td>2.9%</td>
<td>91.4%</td>
</tr>
<tr>
<td>37</td>
<td>2</td>
<td>5.7%</td>
<td>97.1%</td>
</tr>
<tr>
<td>39</td>
<td>1</td>
<td>2.9%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>
"...I found that nearly all the deaths had taken place within a short distance of the pump."

Dr. John Snow, September 1854
Place: Common Exposure Location

• Examples
  – Case 1 – rest. A, B, C, D
  – Case 2 – rest. B, E, F
  – Case 3 – rest. A, B, G, H, I
  – Case 4 – rest. B, D, J, K

Ancel Keys

- **Seven Countries Study**
  - In 1958 professor Ancel Keys, director of the Laboratory of Physiological Hygiene at the University of Minnesota School of Public Health, launched the *Seven Countries Study*. For more than a decade professor Keys, with the help of an international team of specialists, studied the diet, lifestyle and incidence of coronary heart disease among 1middle-aged men from seven countries: the United States, Japan, Italy, Greece, the Netherlands, Finland and Yugoslavia.
Macro Results

Heart disease statistics from the Seven Countries Study

- Crete (Greece): 3
- Ushibuka (Japan): 13
- Montegiorgio (Italy): 24
- United States: 92
- East Finland: 171

Source: www.mediterranian.com/graphics/q&a/ancel.jpg
Water: Another Complex Exposure
Variable- What Has Changed?


Global Microbial Threats in the 1990s

Hantavirus, 1993
Dengue, 1993
Leptospirosis, 1995
Yellow fever, 1995
Cholera, 1991
Cyclospora, 1996
West Nile, 1999
Anthrax, 1993
Dengue, 1994
Ebola, 1994
HIV-1 subtype B, 1994
Bolivian hemorrhagic fever, 1994
Variant CJD Disease, 1996
Diphtheria, 1993
Respiratory Disease, 1993
E. coli O157:H7, 1996
V. cholerae O139, 1992
Plague, 1994
Yellow fever, 1995
Dengue, 1992
Morphivirus, 1994
Assessment of Cause (a life-long exercise!)

- Strength of Association
- Consistency with the existing literature
- Temporal Sequence
- Dose-response present?
- Biological plausibility
- Specificity of the association
- A CAVEAT: this is an art as much as a science
A User Guide for the Epidemiology Literature

- Description of the evidence
- Exposure or intervention
- Outcome
- Study design
- Study population
- Main result
- Non-causal explanations
- Observation bias
- Confounding
- Chance
- Positive features
- Time relationship
- Strength
- Dose-response

- Consistency
- Specificity
- Generalizability
- Eligible population
- Source population
- Other populations
- Comparison with other evidence
- Consistency
- Specificity
- Plausibility
- Coherence
The Good, the Bad, and the Truly Ugly

• The Good:
  – The strengths of epidemiology lie in describing relationships between selected exposures and disease outcomes through the establishment of causality. Also strong for the identification of risk factors.

• The Bad:
  – Complex exposures pose vexing problems.

• The Ugly:
  – Establishing the population-based preventative value of selected interventions (as highlighted by Taubes).
Questions?