DISEASE CLUSTERS, CAUSATION, AND COMMON SENSE

• MOHC Sept, 19-20 2019
• Alan Ducatman, MD, MSc
aducatman@hsc.wvu.edu
Objectives

- Identify three (epidemiologic, historical, and information) paradoxes of disease clustering.
- Obtain a community, epidemiologic, and clinician perspective on the phenomenon of disease clustering, and cost-effective evaluation.
Objectives

• Pinpoint the essential epidemiologic problem (a posteriori logic) in non-epidemiologic terms (submerged comparisons).

• Discuss clusters that led to disease understanding in our lifetimes

• Formulate a cost-effective evaluation approach to cluster events.
Community Questions

• We have too much .............. Didn’t something have to cause this?
• Since it has a cause in the environment, what should we be measuring (biologic monitoring and medical surveillance) to protect ourselves from more of it?
• How do we guarantee (or increase) safety?
Cluster

• “A collection of things of the same kind,” such as grapes – OED

• “A cluster of calamities”
  Edmund Spencer 1590
Disease Cluster

Apparent increases of specific outcome(s) among individuals linked in time/space (cluster) or by exposure characteristics ("aggregation," common in occupational health).

We refer to aggregations as "clusters" in common speech.
Intro to Epidemiology
Paradox

• The study of diseases and disease determinants in populations,
• With statistical assumptions all based on the condition that we do not have advance information and we can test the null hypothesis because..............
• the comparisons are in a priori equal groups.
• Yet, “Clusters” start with unusual numerators and defy the logic.
Cluster Characteristics

- Numerator first
- Rearward projection of linked denominator
- Concerned population, usually
  - Sometimes posed by an astute clinician, and even less commonly by a public health official on whose “turf” the question rests
Epidemiologists’ Questions

• Population and design
• Use of scarce resources. Distraction from “real” duties.
• Is this a “true cluster”? 
• Why do epidemiologists pose these questions? (And, when will we learn better health communications?)
Problems of Cluster Epidemiology

“Texas sharpshooter” (Seymour Grufferman)

- Numerator manipulations and sliding case definitions
- Denominator manipulations in space/time
- Defies null hypothesis
Other, less fundamental Problems

• Case finding Bias: Unequal case finding in index (denominator) and control populations

• And, biostatistical testing when n is small
Back to the fundamental Problem of Cluster Epidemiology

• A more sophisticated analysis—“Multiple submerged comparisons”:
  – A mathematical concept variant of the Texas Sharpshooter
  – What do we mean by that?
Rare Events and Multiple Comparisons

- As populations and connections among populations grow large, rare events
- In cluster comparisons, there are innumerable “hidden” connections that remain invisible until we consider how clusters are made.

Examples: multiple coin flips, lottery outcomes, “birthday paradox.”
Consider the Number of Groups to Which You “Belong”

- Family
- Current Neighborhood, place of origin
- Work colleagues, professional groups, people you serve through your job, counterparts in other states/programs…
- People with whom you went to school, summer camp…
- Faith community and people you know from previous faith communities
Clusters “Happen”

• How many people must a group contain before two people have the same birthday (month and day)?
Submerged comparisons

- Compute probabilities of NO shared birthdays for increasing numbers of people in the group.

- The first person has a 100% unique number (of course)
- The second has a \((1 - 1/365)\) chance (all but 1 number from the 365)
- The third has a \((1 - 2/365)\) chance (all but 2 numbers)
- The 23rd has a \((1 - 22/365)\) (all but 22 numbers). Add these, and its around 50%.

\[
p(n) = 1 - \left( \frac{364}{365} \right)^{n} = 1 - \left( \frac{364}{365} \right)^{n(n-1)/2}
\]
Paradox #1

• Unusual events appear to have a cause.
• Statistically, however, most clusters have no unifying cause other than chance and the necessity that rare events happen with enough comparisons.
“Perfect”  Example:
Larsen, 1956 (‘47)
Wells, 1998 (‘82)
Problems of Cluster Epidemiology

• Human nature and inability to accept:
  – Scientific rejection of population hypotheses
  – Inevitable role of chance
  – Normal uncertainty, cynicism, and doubt
  – Combined History of “power” behavior by authorities
Causal Attributions:

- If it’s good, we caused it ourselves
  - I’m healthy because I work hard at doing what’s right for my body, mind, and spirit.
## What Caused My Child’s Cancer?

N = 500

<table>
<thead>
<tr>
<th>Factor</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental Exposures</td>
<td>303</td>
</tr>
<tr>
<td>Family History</td>
<td>270</td>
</tr>
<tr>
<td>World Dissonance</td>
<td>24</td>
</tr>
<tr>
<td>Cancer Cluster</td>
<td>23</td>
</tr>
<tr>
<td>Stress</td>
<td>22</td>
</tr>
</tbody>
</table>

Ruccione, J Ped Onc Nursing 1990 PMID 2363874
Traditional View of Cluster Epidemiology

- Unacceptable epidemiologic practices
- Implausible etiologic proposals
- No new etiologic agents
- Thus, investigations should be limited
Exceptions to the Rule

• Epidemiologists have recognized the existence of “exceptional” clusters which have some value, usually occupational.

• Example: in the 1980s, Angiosarcoma from vinyl chloride (case report and mention of two other cases) was the standard citation for the rare useful cluster (also a great tox exam question)
However: Occupational Disease Clusters

<table>
<thead>
<tr>
<th>Disease</th>
<th>Agent</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aplastic Anemia</td>
<td>Benzene</td>
<td>1910-1916</td>
</tr>
<tr>
<td>Aplastic Anemia</td>
<td>Radium</td>
<td>1920</td>
</tr>
<tr>
<td>Leukemia</td>
<td>Radium</td>
<td>1925</td>
</tr>
<tr>
<td>Bladder Cancer</td>
<td>Dyes</td>
<td>1933</td>
</tr>
</tbody>
</table>
Clinicians’ Questions

• How do we know what we know? (cause)
• Coherence; i.e., is there a good story? (hypotheses and cause)
• If there is a good story, how would it be retold? (causes are reliable)
# Occupational Disease Clusters

<table>
<thead>
<tr>
<th>Disease</th>
<th>Agent</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mesothelioma</td>
<td>Asbestos</td>
<td>1960-62</td>
</tr>
<tr>
<td>Lung Cancer</td>
<td>BCME</td>
<td>1973</td>
</tr>
<tr>
<td>Angiosarcoma</td>
<td>Vinyl chloride</td>
<td>1974</td>
</tr>
<tr>
<td>Bladder Cancer</td>
<td>MBOCA</td>
<td>1988</td>
</tr>
</tbody>
</table>
Sentinel Occupational Clusters

- Important sources of understanding about (new) etiologic relationships.

- From 1775 to 1991, at least 133 etiologic discoveries derived from sentinel occupational clusters.

- Image from Fleming LE et al, JO(€)M 1991; PMID 1890493
Discoveries continue (bronchiolitis settings since 2007, Jung Lee & Ducatman, ACOEM 2018)

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Population/toxin or allergen</th>
<th># of cases</th>
<th>Author, Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bronchiolitis obliterans</td>
<td>Cookie factory workers/ diacetyl</td>
<td>4 Cases</td>
<td>Cavalcani Zdo R, 2012</td>
</tr>
<tr>
<td>Bronchiolitis obliterans</td>
<td>Coffee-processing facility/ diacetyl</td>
<td>2 Cases</td>
<td>CDC, 2013</td>
</tr>
<tr>
<td>Bronchiolitis obliterans</td>
<td>Flavor manufacturing worker/ diacetyl</td>
<td>2 Cases</td>
<td>CDC, 2007</td>
</tr>
<tr>
<td>Bronchiolitis obliterans</td>
<td>Fiberglass panel manufacturing Workers</td>
<td>6 Cases</td>
<td>Cullinan, Paul, 2013</td>
</tr>
<tr>
<td>Bronchiolitis obliterans</td>
<td>Microwave-popcorn production plant workers/</td>
<td>8 Cases</td>
<td>Kreiss, Kathleen, 2002</td>
</tr>
<tr>
<td></td>
<td>Volatile butter-flavoring ingredients – Diacetyl</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Paradox #2

• Statistically, most clusters have no cause.

• However, much of what we know about etiologic agents has come from clusters.

• And, it is not just occupations
## Community Clusters

<table>
<thead>
<tr>
<th>Disease</th>
<th>Agent</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parkinsonian Syndrome</td>
<td>MPTP</td>
<td>1979</td>
</tr>
<tr>
<td>PCP Pneumonia</td>
<td>HIV</td>
<td>1981</td>
</tr>
<tr>
<td>Eosinophilia Myalgia</td>
<td>“Peak E” tryptophan</td>
<td>1989</td>
</tr>
<tr>
<td></td>
<td>(noted by GP)</td>
<td></td>
</tr>
<tr>
<td>Epilepsy</td>
<td>Cartoons</td>
<td>1997</td>
</tr>
<tr>
<td>Urothelial Cancer</td>
<td>Herbal Medicine</td>
<td>2000</td>
</tr>
</tbody>
</table>

Other: Epping Jaundice, Ginger Jake, Toxic Oil Yusho and Yu-Cheng
What’s real new?

• “Mystery lung illness linked to vaping. Health officials investigating nearly 100 possible cases.” Washington Post 8_16_19
  https://www.washingtonpost.com/health/2019/08/16/mystery-lung-illness-linked-vaping-health-officials-investigating-nearly-possible-cases/?wpisrc=nl_most&wpmm=1
  Surprising? Metals, flavor, nicotine, and partial products of combustion. BTW: reports for 4 years now!!

• Or humidifier lung injury, tracked back to humidifier disinfectants, S. Korea
  https://www.atsjournals.org/doi/10.1513/AnnalsATS.201504-221OC#readcube-epdf
Community Clusters Can Be Infectious

- March 1999, outbreak of encephalitis and pneumonia among 4 workers at a Singapore abattoir.
- Eleven community patients then confirmed to have Nipah virus (previously unknown Hendra-like paramyxovirus), characterized by focal high-intensity lesions on MRI.
- Now an international epidemic, still has occupational/environmental connotations.
Nipah: Another Cluster, Another Community

• Beginning February 1999, three Malaysian pig farmers presented with fever, headache, and altered consciousness. Initial diagnosis was Japanese encephalitis.

• Postmortem CSF showed antibodies against Hendra viral antigens.

• Before epidemic was over, 200 affected, 91 admitted, and 28 died.
Exposures unite diverse appearing risk groups

• When palm sap collectors and those who unfermented drink palm sap began to get sick (primarily in Bangladesh), we learned more about the natural history of exposure
Recent clusters you have may have heard about

Popcorn alveolitis
Indium Lung Disease (alveolar proteinosis)
Aspergillus meningitis (epidural injection)

And some you may not have heard about

- Lychee associated (hypoglycemic) encephalitis
- Aristocholic acid nephropathy (herbs)
Sentinel Occupational Clusters

• Since 1970, average of 2 clusters per year have yielded important new understandings.

• Most common are hypersensitivity and asthma.

• Most interesting may be progressive demyelinating encephalitis in meatpackers (or whatever is going on with our diplomatic staff in Cuba)
Sentinel Health Event (Environmental)

A disease, disability, or untimely death that is environmentally related and whose occurrence may provide the impetus for epidemiologic industrial hygiene survey, material substitution, engineering control, personal protection, or medical care.

Do We Miss Real Causes?

- YES!
- More clusters (and anti-clusters!!) must exist than we ever find.
- Commonly, clusters test the limits of population science because our techniques cannot “prove negatives.”
- Real, unifying causes of clusters are not necessarily implied, because “disproof” is impossible.
Disease Clusters of the Future

• Better population follow-up
• More “positive” evaluations when susceptible populations and intermediate outcomes are considered
• Improved geographic statistical models→ More clusters detected. Is this good?
What Can We Do with a Perceived Cluster?

• Generate and test an etiologic hypothesis
  – Hypothesis must be about exposure and outcome, and it must be testable

• How do we get there?
Responding to Clusters

Recognition and intuition

• Is there a cluster?
• Is it worth pursuing
• Key characteristics of the cluster:
  – Cohesiveness (mechanism for outcome)
  – Biological plausibility
Clinical Analysis

- Case definition:
  - Disease
  - Symptom complex
  - Histopathology

- Latency

- Numerator linkage
Responding to Clusters

Clinical Analysis

• Case verification:
  – Primary (not summary) material
  – Pathologic confirmation preferred for relevant diagnoses
  – Most apparent clusters do not progress beyond this point
Responding to Clusters

Clinical Analysis

• Knowledge base:
  – Literature search
  – Are likely etiologies known?

• Patient characteristics:
  – Pertinent possible etiologies in index cases
  – Inherent risk factors for index cases
  – “Nested case control”
Responding to Clusters

Case Finding

- Index interview
- On-scene professionals
- Questionnaires
- Registries
- Provider data (neighborhood)
- Insurance data
- Public health professionals involved
Responding to Clusters

“Creating” Incidence Data

• “Rearward” denominator projections in time and space, problems and uses

• “Person years” and incidence
Overrated: Statistical Significance

- Presence does not prove biological importance.
- Absence cannot rule out causation, but it does decrease interest.
- A confidence interval can be used to address unusualness of a cluster. A confidence interval cannot be used to prove the presence, or absence, of causation.
More about significance

• The health dept follow-up rarely includes the same numerator as the index cases or the same denominator as the original hypothesis.
• That can be good, or bad.
• May further erode confidence when re-analyses are done.
Responding to Clusters

Interpretation of Biostatistics

• Statistical significance is expected. Presence of “significance” must be interpreted in context.

• Common sense and unifying exposure hypotheses (possible exposure pathway) are more important than statistical significance.
Exposure Pathway. (there is a waste site, everywhere)
Clusters in Context: not just #’s

• Example 1:
  – Three of 25 people living on my block have peripheral neuropathy.

• Example 2:
  – Three of 100 people living in my neighborhood have developed peripheral neuropathy since the batteries were dumped next to the well that is the shared water supply.
Responding to Clusters

Exposure and Outcome Hypotheses

• The reason for the investigation should be stated at the outset to prevent “sharpshooting.”

• The adoption of untestable hypotheses is viewed as concession of defeat.
Other “Issues”

• Involvement of powerful people
  – Politicians
  – Economic interests
• Media coverage
• “Turf conflicts”
• Missing cases, Dueling denominators
• Resources needed to get the job done
Lessons Learned?

Cognitive Development of Yu-Cheng (‘Oil Disease’) Children Prenatally Exposed to Heat-Degraded PCBs

Yung-Cheng Joseph Chen, MD, MPH; Yue-Chang Guo, MD, PhD; Chen-Chin Hsu, MD, PhD; Walter J. Rogan, MD

Objective.—To compare the cognitive development in Taiwanese children who had been exposed prenatally to high levels of heat-degraded polychlorinated biphenyls (PCBs) with control children who were exposed to background levels. This disorder was called Yu-Cheng, “oil disease,” in Taiwan.

Setting.—Communities in central Taiwan in which there had been a cooking oil contamination and mass poisoning by heat-degraded PCBs in 1978 through 1979. An additional community in which these were not present served as a control.

Participants.—One hundred eighteen children born between June 1978 and March 1980 or after their mothers’ consumption of contaminated rice oil; 118 children matched for age, sex, neighborhood, maternal age, and paternal education and occupational class; and 16 older siblings of exposed children, born before the poisoning.

Mein Outcome Measures.—Cognitive development measured from 1985 through 1990 using the Chinese versions of the Stanford-Binet test and the Wachtel Intelligence Scale for Children, Revised.

Results.—The exposed children scored approximately 5 points lower on the Stanford-Binet test at the ages of 4 and 6 years and approximately 3 points lower on the Wachtel Intelligence Scale for Children, Revised, at the ages of 6 and 7 years. Children born up to 6 years after their mothers’ exposure were as affected as children born within a year or two after exposure when examined at 4 and 7 years of age. Older siblings resembled the control children.

Conclusion.—Children prenatally exposed to heat-degraded PCBs had poorer cognitive development than their matched controls. The effect persisted in the children up to the age of 7 years, and children born long after the exposure were still affected.

Polychlorinated biphenyls (PCBs) are a family of industrial compounds that had widespread use from their introduction in the 1930s until their manufacturing ceased in the mid 1970s. Their major use was in electric transformers and capacitors, but the inulating, optical, and thermal properties of PCBs led to their use in cable, carbons, circuit breaker paper, small electric parts, and microwave ovens. Some PCBs contained dioxins, which were discarded without thought to the environmental impact; by 1955 the PCBs had become a worldwide pollution problem, since they resisted physical or biological degradation and accumulated in the food chain. Polychlorinated biphenyls are still routinely detectable in human tissue. Most general population exposure probably occurs through low-level contamination of the food supply, but occupational exposures can produce high body burdens, and special exposures, such as consuming contaminated fish from Lake Michigan, can produce body burdens above background.

Polychlorinated biphenyls, as well as the structurally similar chlorinated naphthalenes and dibenzofurans, had been known from occupational studies to produce a neuropsychiatric syndrome characterized by cognitive impairment. The children of the Yu-Cheng study were born after the poisoning, and although the PCBs were removed from the environment, the children were exposed to PCBs in the womb. This study demonstrates that PCBs can have a deleterious effect on the developing brain at levels of exposure that are not considered toxic. The children were exposed to PCBs in the womb and then to the environment after they were born, and their cognitive development was impaired.

From the Department of Psychiatry, Chiao Tung University, Taiwan, Republic of China, and the National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina.

Reprint requests to Department of Psychiatry, National Cheng Kung University Medical College, Tainan, Taiwan 704, Taiwan, Republic of China.


POLYCLORINATED biphenyls (PCBs) are a family of industrial compounds that had widespread use from their introduction in the 1930s until their manufacturing ceased in the mid 1970s. Their major use was in electric transformers and capacitors, but the insulating, optical, and thermal properties of PCBs led to their use in cable, carbons, circuit breaker paper, small electric parts, and microwave ovens. Some PCBs contained dioxins, which were discarded without thought to the environmental impact; by 1955 the PCBs had become a worldwide pollution problem, since they resisted physical or biological degradation and accumulated in the food chain. Polychlorinated biphenyls are still routinely detectable in human tissue. Most general population exposure probably occurs through low-level contamination of the food supply, but occupational exposures can produce high body burdens, and special exposures, such as consuming contaminated fish from Lake Michigan, can produce body burdens above “background.”

Polychlorinated biphenyls, as well as the structurally similar chlorinated naphthalenes and dibenzofurans, had been known from occupational studies to produce a neuropsychiatric syndrome characterized by cognitive impairment. The children of the Yu-Cheng study were born after the poisoning, and although the PCBs were removed from the environment, the children were exposed to PCBs in the womb. This study demonstrates that PCBs can have a deleterious effect on the developing brain at levels of exposure that are not considered toxic. The children were exposed to PCBs in the womb and then to the environment after they were born, and their cognitive development was impaired.
Responding to Clusters

Communication

• Describe process from the beginning
• Anticipate conflicting agendas
• Maintain open data and recommendations
• Know about available resources
• Do not reach premature closure
Disease Cluster

• An outbreak of disease which all epidemiologists can identify as meaningless. However…..
  – Mesothelioma in asbestos workers (reported but significance missed in 1951)
  – Pneumocystis Carinii pneumonia in Los Angeles (1981) (reported and reinforced with multiple other reports)
Public Health View

• In 1997, state health departments responded to 1100 cluster evaluation requests. Trumbo CW. AJPH 2000;1300-1302
  – Few required further investigation. A way to decide what to evaluate is needed.
  – Yet, this interaction can sometimes represent resources “well spent” (example, bronchiolitis, MPTP)
    • Still a hard call how much to do!
Best Practices: The Clinician’s View

- Open communications
- Public health model
  - conservative
  - inclusive
  - appropriate expertise
    - knowledgeable leader
    - external review if advised
    - no staff vetoes
    - Transparent assumptions and limitations
- Publishable quality and intent
- Recommendations that help the community!!
Paradox #3

• Despite repeated usefulness, there is a scientific prejudice against cluster epidemiology including publication of results.

• As a result, cluster reports often masquerade as (post hoc) epidemiology studies.
Potential Roles of Case Reports and Case Series

- Study of mechanisms
- Medical education and audit
What Agencies and Programs Can Help Communities?

- State registries
- Local and state agencies
- DHHS and DEP
- Federal agencies, roles, and resources
  - ATSDR
  - EPA
- Universities
True vs. False Clusters, health communications 101

• Most clusters are “true” in the sense that they are unusual.
• A few have causes.
• Public health literature distinguishes an approach to “true” (vs false?) clusters, yet this distinction is poor health communication and generally unhelpful, the wrong battle. The key question is – does a cause appear possible? Likely?
Time for an ancient mariner Story?