The take-home message from the online module on the evolution of epidemiology?

The failure to recognize the causes of the plague or to evaluate treatments resulted from a lack of a structured way of thinking about the determinants of disease.
Bruegel’s
*Triumph of Death*
c. 1556
The Black Death of 1349 killed two thirds of Norway’s population.

Cause of the Plague?

- God’s punishment
- Miasmas: unseen vapors from swamps & cesspits
- Contact with lepers
- Walking in the hot sun
The Real Causes:

- *Yersinia pestis*, fleas, and rats
- High population density
- Poor sanitation (garbage attracts rats)

---

Bacteria

Fleas

Sylvatic Cycle

Rodents

Urban Cycle

Garbage

Rats
The inability to identify the determinants & to institute effective preventive measures was not due to a lack of technology.

Why they failed:
- No concept of testing hypotheses in a systematic way in groups of people
- No structured way of evaluating information.

You need to think in a structured way!
A link between antecedent factors and some outcome – possibly a causal relationship, but not necessarily.

**Exposure (Risk Factor)**

Exposures
- “Risk factors”
- Preventive measures
- Management strategy
- Independent variables

**Outcomes**
- Dependent variable
- Disease occurrence

**Examples:**
- Lack of exercise → Heart disease?
- Flu Shot → Dystonia Disorder?
Evolution of Medical Information

1. Descriptive Epidemiology: Hypothesis generation

2. Analytic Epidemiology: Hypothesis testing to establish valid associations

3. Evaluation of efficacy of treatment or prevention: Clinical Trials
Descriptive Epidemiology
Descriptive Epidemiology Provides Clues

What factors might be associated with disease?

- Are there similarities among the diseased?
- Are there differences between diseased & well people?
- What correlates with disease?

- **Person:** characteristics?
- **Place:** specific locations or settings?
- **Time:** does it vary over time?
Identify type & extent of diseases in population. Alert us to new health problems, trends in disease, unusual cases, high risk groups.

Who is getting disease? Their characteristics? (Describe: age, gender, race, geography, habits, diet, drugs used etc.)

How does disease vary across place & time? (trends)

They generate hypotheses for analytic studies.
Sources of Data

- Death Certificates & Birth Certificates; Census
- Disease Registries (cancer, ALS, MS)
- Hospital Discharge Registry
- Infectious Disease Reporting
- Commercial data (sales of tobacco, drugs, etc.)
- Surveys
Large Surveys

- National Survey of Family Growth
- National Health Interview Survey (NHIS)
- National Health & Nutrition Examination Survey (NHANES)
- Behavioral Risk Factor Surveillance System (BFRSS)
- National Health Care Survey
- National Notifiable Disease Surveillance System
- Surveillance of AIDS and HIV Infection
- National Immunization Survey
- Survey of Occupational Injuries and Illnesses
- National Survey on Drug Use and Health
**Differences:** If the frequency of disease differs in two circumstances, it may be due to a factor that differs in the two circumstances. **Example:** stomach cancer in Japan & US

**Similarities:** If a high frequency of disease is found in several different circumstances & one can identify a common factor, then the common factor may be responsible. **Example:** AIDS in IV drug users, or recipients of transfusions, & hemophiliacs.

**Correlations:** If the frequency of disease varies in relation to some factor, then that factor may be a cause of the disease. **Example:** differences in colon cancer vary with per capita meat consumption.
Disease Outbreaks:

How do you know if there is a problem?
COMMUNICABLE AND OTHER INFECTIOUS DISEASES REPORTABLE IN MASSACHUSETTS BY HEALTHCARE PROVIDERS*

*The list of reportable diseases is not limited to those designated below and includes only those which are primarily reportable by clinical providers.

A full list of reportable diseases in Massachusetts is detailed in 105 CMR 300.100.

REPORT IMMEDIATELY BY PHONE!
This includes both suspect and confirmed cases.
All cases should be reported to your local board of health;
if unavailable, call the Massachusetts Department of Public Health:
Telephone: (617) 983-6800  Confidential Fax: (617) 983-6813

REPORT PROMPTLY (WITHIN 1-2 BUSINESS DAYS).
This includes both suspect and confirmed cases.
All cases should be reported to your local board of health;
if unavailable, call the Massachusetts Department of Public Health:
Telephone: (617) 983-6800  Confidential Fax: (617) 983-6813

- Anaplasmosis
- Anthrax
- Any case of an unusual illness thought to have public health implications
- Any cluster/outbreak of illness, including but not limited to foodborne illness
- Botulism
- Brucellosis
- Chagas disease
- Creutzfeldt-Jakob disease (CJD) and variant CJD
- Diphtheria
- Ehrlichiosis
- Encephalitis, any cause
- Food poisoning and toxicity (includes poisoning by ciguatera, scombrotoxin, mushroom toxin, tetrodotoxin, paralytic shellfish and amnesic shellfish)
- Glanders
- Group A streptococcus, invasive
- Haemophilus influenzae, invasive
- Hansen’s disease (leprosy)
- Hantavirus
- Hemolytic uremic syndrome
- Hepatitis A (IgM+ only)
- HBsAg+ pregnant women
- Leptospirosis
- Lyme disease
- Measles
- Melioidosis
- Meningitis, bacterial, community acquired
- Meningitis, viral (aseptic), and other infectious (non-bacterial)
- Meningococcal disease, invasive (Neisseria meningitidis)
- Monkeypox or other orthopox virus
- Mumps
- Pertussis
- Plague
- Polio
- Psittacosis
- Q fever
- Rabies in humans
- Reye syndrome
- Rheumatic fever
- Rickettsialpox
- Rocky Mountain spotted fever
- Rubella
- Severe acute respiratory syndrome (SARS)
- Smallpox
- Tetanus
- Tuberculosis
- Varicella
- Yellow fever

Animal bites should be reported immediately to the designated local authority.
Endemic, Epidemic, Pandemic

Pandemic: Worldwide epidemics

Endemic: Usual occurrence in a geographic area

Epidemic: in excess of normal (1 case of rabies in Newton)
Hepatitis Outbreak

Source?  

20 cases of hepatitis A
Interview Some Cases

Jillian Grey,
Elementary school teacher

Richard Martin,
Father of 8 children,
(5 with hepatitis)

Peter Drew,
Beer distributor

Emily Shultz,
Student & aspiring crime scene investigator
**Based on these clues:**

- Knowledge of biology of hepatitis A (transmission, incubation)
- Time course: epidemic curve of “point source”
- Diverse age, occupation, location
- Interview with a series of cases & similarities in restaurant use

They **hypothesized** that the source was probably an infected food handler at:

- Rick’s Deli
- McDonald’s
- Jaime’s Pub
- Papa Gino’s
- Friendly’s
Characteristics of person, place, & time also generate hypotheses about chronic diseases.
Person: Characteristics of People With Disease

Were they similar with respect to:

- Age, gender, race
- Socioeconomic status
- Body weight
- Physical activity
- Family history
- Diet
- Occupation
- Sexual history
- Travel
<table>
<thead>
<tr>
<th>Age</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-14</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>15-24</td>
<td>2.6</td>
<td>1.6</td>
</tr>
<tr>
<td>25-34</td>
<td>9.4</td>
<td>4.2</td>
</tr>
<tr>
<td>35-44</td>
<td>60.6</td>
<td>16.2</td>
</tr>
<tr>
<td>45-54</td>
<td>265.6</td>
<td>71.2</td>
</tr>
<tr>
<td>55-64</td>
<td>708.7</td>
<td>243.7</td>
</tr>
<tr>
<td>65-74</td>
<td>1670.0</td>
<td>769.4</td>
</tr>
<tr>
<td>75-84</td>
<td>3751.5</td>
<td>2359.0</td>
</tr>
<tr>
<td>85+</td>
<td>8596.0</td>
<td>7215.1</td>
</tr>
</tbody>
</table>
Where Does It Occur?

Does frequency of disease vary with location?

- from country to country?
- from state to state?
- among cities or neighborhoods?
- in different parts of a large workplace?
Stomach Cancer - Location

Females

Males

<table>
<thead>
<tr>
<th>Category</th>
<th>Mortality (/100,000 pop.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese in Japan</td>
<td>58.4</td>
</tr>
<tr>
<td>Japanese immigrants to California</td>
<td>29.9</td>
</tr>
<tr>
<td>Sons of Japanese immigrants</td>
<td>11.7</td>
</tr>
<tr>
<td>Native Californians (Caucasians)</td>
<td>8.0</td>
</tr>
</tbody>
</table>
Does the Rate of Disease Change Over Time?

- Has the frequency of disease changed over several decades?
- Does frequency of disease vary in a cyclic way that relates to the seasons?
- Has it changed over the course of days?
Isolation of exposed people, restricted hospital visits, strict control procedures.

Control precautions re-adopted.

Barrier precautions downgraded.
TB Mortality Over Time

[TB bacillus identified]

[Tuberculin test]

[TB antibiotics]

[BCG vaccine]

Mortality rate per 100,000

Year

1860 1870 1880 1890 1900 1910 1920 1930 1940 1950

Other Factors Can Change Apparent Disease Frequency Over Years

Change in incidence due to environmental or life-style changes. (epidemic of obesity)

Improved diagnosis may increase cases reported even if incidence is not changing. (PSA)

Change in record keeping accuracy

Improved treatment may decrease mortality rates.

Change in the age distribution of a population can produce changes in the overall rate of disease, even though age-specific rates are not changing.
Evolution of Information

Descriptive Studies
- Description; Hypothesis Generation
- Case Report
- Case-Series
- Cross-Sectional
- Correlational

Analytical Studies
- Hypothesis testing
- Case-Control
- Cohort Study
- Clinical Trial
- (Intervention Study)

Compare groups
- Evaluation of Intervention
Descriptive Studies

- Case report
- Case series
- Cross sectional surveys
- Correlational studies

But, can't establish validity of an association.
Case Report & Case Series

Detailed report on one patient (new/unusual).

Or a group with the same problem.

Description only; no hypothesis is being tested.

What factors appear to be associated with development of disease?
AIDS In An Infant: Possible Transmission By Blood Products

- 1983: It was not yet known that AIDS could be transmitted by blood or blood products.

- Infant born with Rh incompatibility; required blood products from 18 donors over 8 weeks.

- Recurrent infections, Candida, decreased T cells.

- No family history of immunodeficiency.

- One of the donors was found to have died of AIDS.
“We believe that AIDS developed in this patient as a result of an infectious agent being transmitted by blood-product administration....”

“Although AIDS as a consequence of a transmissible infectious agent cannot be definitely proven in this patient, the evidence strongly suggests such a possibility. *Future prospective studies* should attempt to determine the incidence of AIDS in transfused patients ....”

Case Series

Pneumocystis Pneumonia in Previously Healthy Young Men

- Previously healthy.
- All had impaired immune function.
- Candida & Cytomegalovirus
- All were sexually active homosexuals

**Hypothesis:** new syndrome of immune dysfunction due to a sexually transmitted agent.
The key to identifying a case series is that the focus is on a single group that is described in detail. Frequently, all of the subjects included in the study have the primary disease or outcome of interest.

For example, an article reported on 239 people who got bird flu. The article might present tables and graphs that gave information about their age, occupation, where they lived, whether they lived or died, etc., but basically it is a detailed description of the characteristics and outcomes in a group of people who all had the same disease.

There is no formal comparison group that was established at the beginning. (They may make some internal comparisons, but the primary goal is to present what happened to a single group.)
Cross-Sectional Surveys

Assess presence of disease & risk factors at a ‘point’ in time

1980          1990          2000          2010

Monitor health status & needs of the population over time. (May also suggest associations between risk factors and diseases).

Example: Health Interview Survey (HIS), a national cross-sectional study for US.

- Current health status
- Habits
- Risk factors
- Demographics
A medical history that you fill out in a doctor’s office is much like a health survey in that it asks about both:

- **Current behaviors (exposures or risk factors)**
  - Do you smoke currently? How much?
  - How many hours per week do you exercise?
  - Did you get a flu shot last year?
  - Do you wear a seatbelt?
  - Do you take vitamin C?

- **Diseases that you have or have had in the past (outcomes)**
  - Have you had a heart attack?
  - Have you been told you have hypertension?
  - Do you have diabetes?
  - Do you suffer from migraine headaches?
  - How much do you weigh? How tall are you?
The key to identifying a cross sectional survey is that information about current health status and current characteristics and behaviors is collected at a single point in time.

These tend to be surveys that ask questions like, “Do you have any of the following diseases?” They also assess current exposure status: Do you smoke, drink, exercise, etc.
Cross Sectional Survey of Heart Disease in Male Farm Owners

<table>
<thead>
<tr>
<th>Status</th>
<th>Prevalence (per 1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not physically active</td>
<td>157</td>
</tr>
<tr>
<td>Physically active</td>
<td>36</td>
</tr>
</tbody>
</table>

What can we conclude?
Cross-sectional surveys ask people their current status with respect to both exposures and diseases. This results in two main disadvantages.

1. The *temporal relationship* between exposure and disease outcomes can be unclear, i.e., which came first.

2. Cross-sectional studies tend to identify prevalent cases of long duration, since people who die quickly or recover quickly or who are no longer employed in a particular occupation are less likely to be identified.
Do you have…?
Are you active?

Which Came First?

Heart
Inactive? → Disease

or

Heart
Disease → Inactive?
Sometimes Cross-Sectional Studies Can Be Analytical

However…

Salary of Assistant Professors

\[
\begin{array}{c|c|c|c|c|}
& > \$60,000 & <\$60,000 \\
\hline
Male & 122 & 75 \\
\hline
Female & 64 & 50 \\
\end{array}
\]

Here, the exposure (gender) clearly was established before the outcome (salary), i.e. the temporal relationship is clear.
Is HIV Transmitted by Insects?

Survey Questions:
• Are you HIV+?
• Do you have any of the following exposures?

<table>
<thead>
<tr>
<th>Exposure</th>
<th>HIV-seropositive Yes/total</th>
<th>HIV-seronegative Yes/total</th>
<th>Odds Ratio</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mosquito bites?</td>
<td>9/56 (16)</td>
<td>97/800 (12)</td>
<td>1.39</td>
<td>0.61-3.05</td>
<td>0.39</td>
</tr>
<tr>
<td>Lice in house?</td>
<td>1/54 (2)</td>
<td>18/811 (2)</td>
<td>0.83</td>
<td>0.02-5.47</td>
<td>0.86</td>
</tr>
<tr>
<td>Mites in house?</td>
<td>1/53 (2)</td>
<td>11/809 (2)</td>
<td>1.40</td>
<td>0.03-9.94</td>
<td>0.75</td>
</tr>
<tr>
<td>Fleas in house?</td>
<td>4/55 (7)</td>
<td>74/808 (9)</td>
<td>0.78</td>
<td>0.20-2.21</td>
<td>0.64</td>
</tr>
<tr>
<td>Bedbugs?</td>
<td>1/56 (2)</td>
<td>8/812 (1)</td>
<td>1.83</td>
<td>0.04-14.0</td>
<td>0.57</td>
</tr>
<tr>
<td>No Insect repellent?</td>
<td>28/44 (64)</td>
<td>316/461 (69)</td>
<td>0.92</td>
<td>0.45-1.90</td>
<td>0.80</td>
</tr>
<tr>
<td>Dengue antibodies?</td>
<td>12/52 (23)</td>
<td>119/798 (15)</td>
<td>1.71</td>
<td>0.82-3.50</td>
<td>0.11</td>
</tr>
<tr>
<td>SLE antibodies?</td>
<td>6/42 (14)</td>
<td>101/798 (13)</td>
<td>1.15</td>
<td>0.42-2.14</td>
<td>0.76</td>
</tr>
<tr>
<td>History of STD*</td>
<td>10/35 (29)</td>
<td>28/425 (7)</td>
<td>5.67</td>
<td>2.28-13.9</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Syphilis antibodies*</td>
<td>14/33 (42)</td>
<td>98/425 (23)</td>
<td>2.46</td>
<td>1.12-5.37</td>
<td>0.01</td>
</tr>
</tbody>
</table>

*Adults only

Science 1988; 239: 193-7
Cross-sectional surveys ask people their current status with respect to both exposures and diseases. This results in two main disadvantages.

1. The *temporal relationship* between exposure and disease outcomes can be unclear, i.e., which came first.

2. Cross-sectional studies *tend to identify prevalent cases of long duration*, since people who die quickly or recover quickly or who are no longer employed in a particular occupation are less likely to be identified.
**Example:** Isocyanates can cause occupational asthma. Continued exposure after development of occupational asthma is associated with developing more severe disease.

In a cross-sectional study of isocyanates and occupational asthma, the prevalence of asthma was lower in factory workers with >5 years employment vs. those with <5 years employment, because those with isocyanate exposure and asthma were more likely to leave.
A Correlational Study (Ecologic Study)

22 countries (populations)

Use multiple groups or populations

Average Meat consumption (in many people)
Advantages:

- Data sets are readily available: quick & inexpensive.
- The correlation coefficient ("r" value) gives a measure of how close the points are to a straight line.

Perfect + correlation: $r = 1.0$

Perfect - correlation: $r = -1.0$

No correlation: $r = 0$

$r = 0.54$

$r = -0.86$

FYI: The EpiTools.XLS spreadsheet has a worksheet that shows how to calculate correlation coefficients using Excel.
Limitations:

- Exposure is measured as the **average for a population**, not a person, so there is **no real link** between exposure disease. (the ecologic fallacy)
- Can’t adjust for other factors affecting outcome (confounding).
- A correlation doesn’t establish causality.
- Complex relationships can be masked.
The key to identifying a correlational study (ecologic study) there is **no information about individual people!!**

It is all based on *average exposures* in multiple *groups* of people.
Another limitation of correlational studies is that average exposures can mask non-linear relationships between exposure and outcome.
This comparison of disease in populations suggests an inverse (negative) correlation.
If we look at *individuals*, there is actually a “J”-shaped relationship between alcohol consumption & CHD mortality.

The graph below is from a cohort study with exposure and outcome data on many *individual* people; it is not a “correlational study”, which only has data on whole groups.

Non-linear relationships like this may be masked by correlational studies.
Descriptive Studies - Summary

- Case report
- Case series
- Cross sectional surveys
- Correlational studies

• Focus is on description: age, gender, race, geography, habits, diet, drugs used etc.

• Alert medical community to new health problems or unusual cases.

• Generate hypotheses.

Something to think about: If the cheerleader on Inside Edition had really had dystonia, what descriptive epidemiology studies might have provided clues about the cause?