What to do today (Tuesday Jan 10, 2023)?

Part I - Introduction

I.1 General Introduction

I.2 A Classic Epidemic Example

Part II - Epidemiologic Methods

Part II.1 Epidemiologic view of diseases and populations

Part II.2 Measuring disease frequency in population

Part II.3 Design of medical studies: cohort and case-control studies; controlled clinical experiments

I.2. An Epidemic of Blindness in Young Children

(Koepsell and Weiss, 2003, Chp1)

- **Discovery**
- **Epidemic**
- **Early Search for Possible Causes**
- **Early Treatments**
- **Narrowing the Search for Causes**
- **Implicating Oxygen: Experimental Evidence**
- **Declining Incidence**
- **Mechanism**
- **Epilogue**
- **Summary**

Part I.2 An Epidemic Example: Blindness in Young Children (Koepsell and Weiss, 2003, Chp1)

Part I.2.1 Discovery

- On Feb 14, 1941, a Boston pediatrician Dr. Stewart Clifford was puzzled and concerned by a rountine house call on a baby girl.
- Dr. Paul Chandler, a leading Boston ophthalmologist, was called.
- Within a week, Dr. Clifford encountered another blind infant with the condition.
- Dr. Theodore Terry, a consultant ophthalmologist, had collected information on five such cases in the Boston area. (Terry, 1942, *Amer J of Ophthalmology*)
- The shared pathological features among these early cases: *retrolental fibroplasia* (RLF)

Part I.2.2 Epidemic

After Terry's description of RLF appeared in the medical literature,

- In 1945, Terry himself reported on 117 cases: all but five of them babies born prematurely
- The California School for the Blind found a sharp rise in the number of RLF:
  - 1945: 2; 1946: 8; ... 1950: 61; 1951: 88
- During the decade after 1941, RLF went to being the most common cause of blindness in preschool children in US. World wide, $> 10,000$ babies developed the condition.

Part I.2.3 Early Search for Possible Causes

The **rapid rise in the frequency** of RLF and its poor prognosis led to an intensive search for its cause or causes.

- Early cases suggested that most cases had been born prematurely.
- The care of premature infants had advanced in many ways during 1940s.
  - overall survival of premature babies $\nearrow$
  - # of RLF $\nearrow$
  - the proportion of survivors' RLF $\nearrow$ remarkedly

Two possibilities:

- RLF is a complication of prematurity;
- RLF is related to techniques used to support premature infants

Part I.2.4 Early Treatments

A group of New York-based physicians tried ACTH treatments for babies with early signs of RLF:

- 25 out of 31 appeared to respond: early reports of success with ACTH was welcomed news (Blodi et al, 1951)
- noted many disturbing treatment failures (Laupus, 1951; Pratt, 1951)

An attempt to provide more convincing evidence:

- Reese et al (1952) undertook a second study of ACTH treatment.

$\Longrightarrow$ **A rude shock:** ACTH had been found to be ineffective and even dangerous as treatment for RLF, and quickly fell out of favor.

Part I.2.5 Narrowing the Search for Causes

A decade after the 1st RLF case, over 50 factors had been suggested as possible contributors ...

- excluding a few that had changed little during the period of rising RLF, still a long remaining list.
- Kinsey and Zacharias (1949) reported 3 related studies
  - Study 1. all babies with weight < 4 at birth, born at a Boston hospital over 10 years, were identified, and compaired between 53 babies developed RLF (cases) and 298 babies no-RLF (controls)
    $\Longrightarrow$ 3 factors associated with differences larger than chance alone could easily explain
  - Study 2. the number of RLF cases $\nearrow$: whether changes over time in use of any particular treatment coincided with the frequency of RLF change
    $\Longrightarrow$ suspicious (i) use of iron supplements, (ii) use of water-miscible vitamins, (iii) amount of supplemental oxygen used

Part I.2.5 Narrowing the Search for Causes

- Kinsey and Zacharias (1949) reported 3 related studies
  - Study 3. incidence of RLF seemed to differ among hospitals, and hospitals had different policies and practices for treating premature babies
    $\Longrightarrow$ hospital-to-hospital variation in RLF frequency and in treatment might be linked

- Dr. Kate Campbell, an Australian pediatrician, conducted a small study (1951)
  - compared premature babies treated at three institutes: Institute I with oxygen at high concentration; Institutes II and III with low – 19% of RLF at Institute I vs 7% at Institutes II and III

$\Longrightarrow$ oxygen supplementation might actually be causing RLF?

Part I.2.6 Implicating Oxygen: Experimental Evidence

Dr. Arnall Patz, then an ophthalmology resident at a D.C. hosipital

- began with some animal studies and reasoned that more direct and convincing evidence about oxygen supplementation was a cause of RLF
- obtained $4000 grant from NIH for a study of human babies
  - Infants < 3.5 pounds were assigned alternately to receive oxygen concentrations of:
    (i) 65 - 70% (high-oxygen group) for four to seven weeks;
    (ii) $< 40\%$ (restricted-oxygen group) only in response to clinical need and only from one to fourteen days
  - 11/76 were excluded; 17/28 (61%) and 6/37 (16%) of RLF in (i) and (ii), respectively
  - a difficult study to carry out: nurses questioned the wisdom of curtailing oxygen and turned up the oxygen concentration at night for some babies in (ii)

Part I.2.6 Implicating Oxygen: Experimental Evidence

National Cooperative Study (Kinsey, 1955)

- a randomized clinical trial in 18 hospitals using a common protocol, i.e. a multicenter clinical trial

Treatment	Mortality death/N (%)	RLF Incidence cases/N (%)
routine high oxygen	15/68 (22%)	12/53 (22.6%)
curtailed oxygen	36/144 (25%)	8/104 (7.7%)

$\Longrightarrow$ total mortality in 2 groups similar; RLF 3-fold higher in high-oxygen group

A smaller randomized trial in Colorado reached similar conclusions.

Part I.2.7 Declining Incidence

After the release of the NIH study findings, the American Academy of Pediatrics and other influential professional organizations soon revised their recommendations to clinicians on care of premature babies, advocating more sparing use of oxygen ...

The number of RLF cases in USA fell rapidly ...

- In Southern California, 1951: 88; 1952: 83; ... 1955: 9; 1956: 2; 1957: 1 ...

$\Longrightarrow$ The epidemic was over.

Part I.2.8 Mechanism

- Oxygen was obviously essential for human life
- Premature babies oftern suffered problems that led to low blood oxygen levels

why was more of a good thing not better?

- Ashton et al (1953) reported research findings from work on kittens and explained it.

Part I.2.9 Epilogue

- The rapid shift to more conservative use of oxygen supplementation in 1950s
  - generally credited as being the main cause of the reapid decline in RLF
  - with cost!
    e.g. reported deaths had increased in the period at Johns Hopkins Hospital: Some clinicians may have been overzealous in their efforts to prevent RLF.
  - small to moderate adverse effects of oxygen curtailment on mortality?

- Clinicians are far better able to monitor and regulate oxygen levels now.

Part I.2.10 Conclusions

Recount the history of retrolental fibroplasia (RLF) ...

- the value of information gained from studying variations in disease frequency in human populations
  - the center stage of the epidemiologic research: enabling rapid recognition of an important cause of disease

- few of the investigators regarded them as epidemiologists: then most epidemiologists worked on infectious diseases
  $\Longrightarrow$ other types of epidemiologic studies

Part I.2.10 Conclusions

- features of a research design make one study better than another, affect our confidence in the validity of the results
  - e.g. early enthusiasm for ACTH for RLF due to treated cases, no comprable group of cases treated without ACTH
  - e.g. Owens and Owens' *prospective cohort* study to confirm the role of prematurity
  - e.g. Kinsey and Zacharias' *case-control* study to investigate various risk factors, their *ecological* research designs
  - e.g. Patz's *non-randomized intervention trial*
  - e.g. the National Co-operative Study, a true *randomized controlled trial*

$\Longrightarrow$ **Have you been motivated enough to move on?**

Part II - Epidemiologic Methods

Part II.1 Epidemiologic view of diseases and populations
Example. Crohn's Disease - Manitoba, 1989-1994
(Crohn's disease is an inflammatory bowel disease causing pain, diarrhea and blood loss.)

- Bernstein et al. (1999) estimated the proportion of Manitoba residents with Crohn's disease, 1989-1995.
- Bernstein et al. (1999) used the following case definition: "individuals registered with Manitoba Health for at least 2 years between 1984 and 1995 were classified as having Crohn's disease or ulcerative colitis only if they had at least five separate medical contacts with such a diagnosis. Individuals who were registered for less than 2 years during the study period were classified as cases if they had had at least three separate medical contacts."

the same as the disease diagnosis?

Part II.1.1 Case Definition vs. Diagnosis

- A case definition is an operational definition of a disease for the purpose of a specific study.
- Case definitions usually differ from diagnoses:
  - Diagnoses direct treatment and suggest prognoses;
  - Case definitions are used, for example, to measure the burden of disease in a population.
- Case definitions may be broader or narrower than diagnoses, e.g.
  - Motor vehicle crash injuries,
  - Cases in the Manitoba Crohn's Disease study

Part II.1.2 Reasons for Between Study Variation in Case Definitions

Difficulties in applying clinical diagnosis

- Changes in clinical diagnosis over time:

CDC definition of AIDS since 1981 Case Report
(Aschengrau and Seage, 2008 Table 2-2)

Year	AIDS Science?	Clinical diagnosis
1982	Very limited	Kaposi's sarcoma, Pneumocystis pneumonia, other severe opportunistic infections
1985	HIV discovered as cause, antibody test developed	23 clinical conditions & lab evidence of infection
1993	Discovered role of T cells in disease progression	26 clinical conditions & symptom free with low CD4 T cell counts

In 1993 the number of U.S. AIDS cases increased by 75%.

Model of Disease Natural History

Biological onset	Detectable by testing	Symptoms begin	Diag -nosed	Becomes disabling

Time

This model explains temporal changes in the AIDS case definition

Part II.1.3 Reasons for Between Study Variation in Case Definitions

Other reasons ...

- Absence of a common clinical diagnosis, e.g. the international classification of ROP was only introduced in 1984 (Quinn, 2005).
- Broad vs. narrow case definition: the primary interest?
- Source of data: limitations

Part II.1.4 Implications of between study differences in case definition?

- On estimates of disease frequency
- On estimates of exposure effects

$\Longrightarrow$ *Be Careful with Interpreting the Analysis Results*

Part II.1.5 Generic Disease Model

Susceptible Diseased

Not at Risk

- Prevalence = distribution of population among boxes at a point in time.
- Incidence = flow rate from susceptible to diseased.

Disease onset as an abrupt transition
Time points used to determine start of disease?

Biological onset	Detectable by testing	Symptoms begin	Diag -nosed	Becomes disabling

Time

Part II.1.6 Defined Populations

For example, ...

- Bernstein et al. (1999) estimated the proportion of Manitoba residents living with Crohn's disease, 1989-1995.

- To do so they needed to define what it means to be a Manitoba resident.
  - A population must be defined before its size can be determined.

  - Define population by specifying characteristics unique to members.

  - Manitoba residents were defined as people who held Manitoba Health Cards between 1989 and 1995.

  - Advantages /disadvantages of this definition?

Part II.1.7 Link Between Cases and Population at Risk

- Why is it not necessary to identify all cases in a defined population when estimating the proportion of people with the disease?

- When do clinical cases correspond to an identifiable population at risk?
  - Population-wide case registry
  - Health care for members of a single health insurance plan

- What if defined population observed over time?
  - Open populations: may gain and lose members over time (e.g. Manitoba population 1989-1995).
  - Closed populations: gains no new members (once complete) and loses members only to the disease of interest.

Part II.2.1 Review of the Basic Concepts

Rothman (2012): **epidemiology** is "the study of the occurrence of illness"
$\Longrightarrow$ *epidemiologic studies are to identify all cases of a disease in a defined population at risk* and analyze them
(**"disease" or "event", a generic term**)

case definition: the operational definition of a disease for study purposes

- epidemiologic case definition vs clinical diagnosis
- disease models: disease states – "non-diseased", "diseased"; related times – "onset" time, "death" time
- *susceptible* (i.e. at risk) – capable of becoming a case of the disease

recurrent event, *abrupt event*

Part II.2.1 Review of the Basic Concepts

Relating cases to the base population: central to epidemiologic thinking

$\Longrightarrow$

population: to specify the characteristics of its members in common

- closed vs open populations
- defined populations
- link between cases and the population at risk: to identify all cases

Part II.2.2 Disease Frequency: Descriptive

- **Why are case counts often insufficient?**
  geographic and temporal comparisons
  $\Longrightarrow$ numerator and denominator data?
- **To choose an appropriate measure for a certain purpose**
  - What kind of information is needed?
  - What kind of information is obtainable?

Examples:

- for the BC goverment to budget for covering the cost to treat HIV+ patients in 2012
- for the BC goverment to find out the change in HIV+ in 2012 from 2011

Part II.2.2 Disease Frequency: Descriptive

What kind of information is needed?

- to know how *common* (or wide spread) a disease is
- to know how *often* new cases of disease develop

$\Longrightarrow$ **prevalence** vs **incidence**

- **prevalence**: the proportion of a population found to have a disease

$$prevalence = \frac{NumPrevalentCases}{SizePopulation}$$

- **incidence**: a measure of the risk of developing some new cases of the disease within a specified period of time

$$incidence = \frac{NumIncidentCases}{NumPeopleAtRisk}$$

Part II.2.2 Disease Frequency: Descriptive

Prevalence (or Prevalence Proportion):

- "Point prevalence": at a specific point in time.

- "Period prevalence": during a given period (e.g. one-year prevalence, etc)
  including people who already have the condition at the start of the study period as well as those who acquire it during that period

- "Lifetime prevalence": the proportion of a population that at some point in their life (up to the time of assessment) have experienced the condition

Limitations?

Part II.2.2 Disease Frequency: Descriptive

Incidence: (better expressed as a proportion or a rate with a denominator)

- "incidence proportion", "cumulative incidence"
- "incidence rate": the number of new cases per population in a given time period
  - in practice, often assumed implicitly the incidence rate is constant over different periods of time (or the average of the rates during different time periods)

Part II.2.2 Disease Frequency: Descriptive

Incidence vs Prevalence?

Example. a disease that takes a long time to cure and was widespread in 2002 but dissipated in 2003:

- both high incidence and high prevalence in 2002
- in 2003 it will have a low incidence yet will continue to have a high prevalence

compared to a disease that has a short duration may have a low prevalence and a high incidence

Practically **prevalence = incidence * duration**.

Part II.2.2 Disease Frequency: Descriptive

Example.

- Bernstein et al. (1999) reported that in Manitoba as of Dec. 31, 1994, Crohn's disease prevalence was

$$\frac{2268}{1140000} = 199 cases/100K$$

using provincial health care records.

- Bernstein et al. (1999) identified 997 incident cases of Crohn's disease, 1989-1994.
In 1991 there were 1,091,942 people living in Manitoba.
Crohn's disease incidence rate in Manitoba, 1989-1994:

$$\frac{997}{6 \times 1091942} = 0.000152,$$

15.2 cases per 100K person-years.

Part II.2.2 Disease Frequency: Descriptive

Period prevalence combines information about point prevalence and cumulative incidence

Example. What do they tell about?
Among N women giving birth to live born children

- D=# with diabetes = d+g
  - d=# with diabetes at conception
  - g=# developing diabetes while pregnant
- Point Prevalence = d/N
- cumulative incidence= g/(N-d)
- Period Prevalence = D/N

$$\frac{d+g}{N} = \frac{d}{N} + \frac{g}{N-d} \times \frac{N-d}{N}$$

What to study next?

Part I – Introduction

Part II - Epidemiologic Methods

- *II.1 Epidemiologic view of diseases and populations*
- II.2 Measuring disease frequency in population
- II.3 Design of medical studies: cohort and case-control studies; controlled clinical experiments

Part III - Clinical Trials

Part IV - Mordern Analytic Approaches