Evaluating Screening Programs & Health Interventions

David Celentano, ScD, MHS
Learning Objectives

- Review criteria to evaluate screening effectiveness
- Describe the natural history of disease and its impact on screening
- Review epidemiologic designs to evaluate:
  - Screening programs and outcomes
  - Health services and clinical interventions
- Evaluate sources of bias in screening
- Describe how measurement of a health intervention can affect how we determine effectiveness
Epidemiology in Public Health Practice

Evaluating Screening Programs
Accounting for the Natural History of Disease in Screening
Hepatitis C infection $\rightarrow$ YEARS

- **Preclinical Phase**
- **Clinical Phase**

Biologic Onset of Disease $\rightarrow$ Symptoms $\rightarrow$ Diagnosis $\rightarrow$ Therapy $\rightarrow$ Outcome

Influenza infection $\rightarrow$ hours/days
Preclinical Phase

Biologic Onset

Symptoms

Disease Detectable by Screening

Clinical Phase

Diagnosis

Therapy

Outcome
An example from HIV/AIDS

Natural History and Laboratory Staging of HIV Infection

Eclipse Phase

V RNA+
p24Ag+
ELISA+
Western blot +/−
Western blot + (p31−)

Plasma virus RNA (copies/ml)

Viral RNA cutoff 50 copies/ml
Ultrasensitive Viral RNA cutoff 1-5 copies/ml

Days following HIV-1 transmission

(adapted from Fiebig, AIDS 2003)
Possible Outcomes in Epidemiologic Studies to Evaluate Screening Programs
Assessing the Effectiveness of Screening Programs Using Operational (process) Measures

- Number of people screened, and proportion of target population screened
- Detected prevalence of preclinical disease
- Costs per case found
- Proportion of positive persons brought to final diagnosis and treatment
- Predictive value of a positive test in population screened

Adapted from Hulka BS. Cancer 1988;62:1776-80
Assessing the Effectiveness of Screening Programs Using Outcome Measures

1. Reduction of mortality in the population screened
2. Increase in percent of cases detected at earlier stages
3. Reduction in complications
4. Prevention of recurrence
5. Improvement of QoL in screened individuals
Possible Outcomes of a Screening Program

Usual Diagnosis

Earlier Diagnosis

Usual Death
Possible Outcomes of a Screening Program

Usual Diagnosis

Usual Death

Earlier Diagnosis

Usual Death

Earlier Diagnosis

Delayed Death
Possible Outcomes of a Screening Program

Usual Diagnosis

Earlier Diagnosis

Usual Death

Usual Diagnosis

Usual Death

Earlier Diagnosis

Delayed Death

Earlier Diagnosis

No Deaths
Assumptions Underlying a Relationship of Improved Outcome to Early Detection of Disease

1. All or most clinical cases of a disease first go through a detectable preclinical phase

2. In the absence of intervention, all or most cases in a preclinical phase progress to a clinical phase
Assumptions Underlying a Relationship of Improved Outcome to Early Detection of Disease

1. All or most clinical cases of a disease first go through a detectable preclinical phase

2. In the absence of intervention, all or most cases in a preclinical phase progress to a clinical phase
Natural History of Cervical Cancer - I

NORMAL \[\rightarrow\] CARCINOMA IN SITU \[\rightarrow\] INVASIVE CANCER

(months) (years)
Natural History of Cervical Cancer - II

NORMAL \rightarrow CARCINOMA IN SITU \rightarrow INVASIVE CANCER

\rightarrow
Possible Study Designs to Evaluate Screening Programs
Design of a **Non-Randomized** Comparison (Cohort) Study of the Benefits of Screening

- **Screened**
  - Die from the Disease
  - Don’t Die from the Disease

- **Not Screened**
  - Die from the Disease
  - Don’t Die from the Disease
1. Selection bias
   a. Referral (Volunteer) bias
1. Selection bias
   a. Referral (volunteer) bias
   b. Length-biased sampling
      A systematic error due to selection of disproportionate numbers of long-duration cases (long-term survivors) in one group.
Short Natural History:

Preclinical Phase | Clinical Phase

Long Natural History:

Preclinical Phase | Clinical Phase
Problems Which Complicate Assessment of Improvement in Survival as a Result of Early Detection

1. Selection bias
   a. Referral (volunteer) bias
   b. Length-biased sampling

2. Lead-time bias
Lead Time

- Lead time

  interval by which the time of diagnosis can be advanced by the screening procedure, as compared with the normal methods for detection and diagnosis
Five-Year Survival and Lead-Time Bias I

Biological Onset of Disease

Diagnosis and Treatment

Death

SURVIVAL

1998

2011

2019
Five-Year Survival and Lead-Time Bias II

- Biological Onset of Disease
- Usual time of Diagnosis and Treatment
- Onset
- Detected by Screening: Diagnosis and Treatment
- SURVIVAL
- 1998
- 2005
- 2011
- 2019
- Death
LEAD TIME BIAS – I
5-Year Survival When Diagnosis is Made Without Screening

![Graph showing cumulative survival rates over 5 years after diagnosis. The graph indicates that 30% of diagnosed cases survive the first 5 years, with a steep decline in survival thereafter. A red box highlights the 5-year survival period.](image)
LEAD TIME BIAS – II
Shift of 5-Year Period By Screening And Early Detection
(Lead Time)
LEAD TIME BIAS – III
Bias In Survival Calculation Resulting From Early Detection
Design of a Randomized Trial of the Benefits of Screening: HIP Study

HIP Enrollees
~ 62,000

Randomized

Screening Including Mammography
~31,000

Breast Cancer

No Breast Cancer

Regular Care
~31,000

Breast Cancer

No Breast Cancer

Compare Mortality
Deaths Due to Breast Cancer
HIP 5-Year Follow-Up

Numbers

Rates

Number of Deaths:
- Control: 63
- Study: 40

Death Rate per 10,000 Person-years:
- Control: 4.1
- Study: 2.6
Case Fatality due to Breast Cancer
HIP 5-Year Follow-up

<table>
<thead>
<tr>
<th>Study Group</th>
<th>Case Fatality Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Group</td>
<td>40%</td>
</tr>
<tr>
<td>Total Study</td>
<td>29%</td>
</tr>
<tr>
<td>Refused Screening</td>
<td>35%</td>
</tr>
<tr>
<td>Total Screened</td>
<td>23%</td>
</tr>
<tr>
<td>Detection Due to Screening</td>
<td>13%</td>
</tr>
<tr>
<td>Detection NOT Due to Screening</td>
<td>38%</td>
</tr>
</tbody>
</table>
Thought experiment

- How would you interpret results that show no benefit from screening?
Interpreting Results of No Benefit of Screening

- The apparent lack of benefit may be inherent in the *natural history of the disease*

- The therapeutic intervention *may not be any more effective when it is provided earlier* in the natural history of disease

- *Inadequacies in the care provided* may account for the observed lack of benefit
Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness

H. Gilbert Welch, M.D., M.P.H., Philip C. Prorok, Ph.D., A. James O’Malley, Ph.D., and Barnett S. Kramer, M.D., M.P.H.
**Figure 1.** Temporal Relationship between the Introduction of Screening Mammography and Increased Incidence of Invasive Breast Cancer.
CONCLUSIONS
Although the rate of detection of large tumors fell after the introduction of screening mammography, the more favorable size distribution was primarily the result of the additional detection of small tumors. Women were more likely to have breast cancer that was overdiagnosed than to have earlier detection of a tumor that was destined to become large. The reduction in breast cancer mortality after the implementation of screening mammography was predominantly the result of improved systemic therapy.
Epidemiology in Public Health Practice

Evaluating Health Interventions
Efficacy

The extent to which a specific intervention or service produces a beneficial result under ideal conditions (based on results of an RCT)
Effectiveness

The extent to which a specific intervention or service, as deployed in the field, does what it is intended to for a defined population.
Efficiency

The effects or end-results achieved in relation to the effort expended – cost/benefit (in terms of money, resources and time, a measure of economy with which a procedure of known efficacy and effectiveness is carried out)
Problems in Comparing the Effects of Medical Care in Two Population Groups

- Are the characteristics of the populations comparable?
  - demographically and in factors relating to risk and prognosis

- Are the measurement methods comparable?
  - diagnostic methods and classification
Outcomes Research
Some Health Endpoints Used in Outcomes Research

- Morbidity
- Mortality
- Quality of life (Euro-QAL)
- Functional status (ADL, IADL)
- Patients’ perceptions of health status
  - Symptom recognition
  - Patient satisfaction and pain levels
It seems to work in clinical investigation...
Some Economic Endpoints Used in Outcomes Research

- Hospitalization rates, readmissions within 30 days
- Outpatient (+) and emergency room visits (-)
- Lost days of work
- Days of restricted activity like DALY
- Out-of-pocket expenses
Variation in Surgical-Readmission Rates and Quality of Hospital Care

Thomas C. Tsai, M.D., M.P.H., Karen E. Joynt, M.D., M.P.H., E. John Orav, Ph.D., Atul A. Gawande, M.D., M.P.H., and Ashish K. Jha, M.D., M.P.H.

Readmission 12.7% vs. 16.8% for high vs. low volume quartiles
Possible Study Designs to Evaluate Health Interventions
Cohort
1. Randomized
Effect of Clinical Geriatric Assessments and Collaborative Medication Reviews by Geriatrician and Family Physician for Improving Health-Related Quality of Life in Home-Dwelling Older Patients Receiving Polypharmacy: A Cluster Randomized Clinical Trial

Rita Romskaug, MD; Eva Skovlund, MSc, PhD; Jørund Straand, MD, PhD; Espen Molden, MSc, PhD; Hege Kersten, MSc, PhD; Kaisu H. Pitkala, MD, PhD; Christofer Lundqvist, MD, PhD; Torgeir B. Wyller, MD, PhD
Figure 2. Primary Outcome of Health-Related Quality of Life as Measured by the 15D Instrument

Shown are mean (SD) 15D instrument scores at baseline, week 16, and week 24. The score range is 0 to 1, with higher scores indicating better quality of life.
Types of Epidemiologic Study Designs for Evaluating Health Services

- **Cohort**
  1. Randomized
  2. Non-randomized

Watch out for bias!
Types of Epidemiologic Study Designs for Evaluating Health Services

- Cohort
  1. Randomized
  2. Non-randomized
     ▶ Before/after
Cases (n=300) severe rotavirus gastroenteritis cases in Nicaragua, 792 hospital controls, 851 community controls. Vaccine coverage of RV5 reached 92% VE = 87% for community controls receiving 3 doses followed two years, 64% in hospital controls VE = 85% in children <12 months at RGE onset
DM2 cases ages 15-34 registered in Sweden in 1983 matched with two population controls
1991 interviewed on retrospective data on utilization in prior three months
74% of cases and 19% of controls reported at least one OPD hospital visit
- OR=14 for one visit
- OR=11 for two visits
- OR=8.9 for 3+ visits

Deciding Program and Policy Priorities

Effectiveness of Prevention

Importance of the Problem in the U.S.

Low

High

CHD

Low

High

61
Deciding Program and Policy Priorities

Effectiveness Of Prevention

CHD

High

Low

Importance of the Problem in the U.S.

Low

High

Rheumatic Fever
Studies of Disparities in Health Care
Racial Disparities in the US

- Asthma prevalence elevated in Puerto Ricans and African Americans
- HIV/AIDS 8 times more common in African Americans than Whites
- African Americans at greater risk for ESRD, donate organs proportionately, but more often cannot get transplants
- African Americans account for ~13% of US population but 70% of gonorrhea and half of syphilis and chlamydia
EFFECTS OF RACE AND INCOME ON MORTALITY AND USE OF SERVICES AMONG MEDICARE BENEFICIARIES

MARIAN E. GORNICK, M.S., PAUL W. EGERS, PH.D., THOMAS W. REILLY, PH.D., RENEE M. MENTNECH, M.S., LESLYE K. FITTERMAN, PH.D., LAWRENCE E. KUCKEN, M.P.A., AND BRUCE C. VLADECK, PH.D.

NEJM 1996;335:791
Thought Experiment 2

If we control for income, are racial disparities in mortality reduced?

If so, why?

If not, why not?
Rates of Mortality By Race, Sex and Income Among Medicare Beneficiaries 65 and Older

Gornick et al, 1996
Advantages of Using Large Data Sets

- Data refer to real-world populations, so the problem of external validity is minimized
- Analysis can usually be completed quickly because the data have already been collected
- Sample size is usually not a problem except in specific subgroups
Disadvantages of Large Data Sets - I

- Since data were gathered for fiscal or administrative purposes, they are often not well suited for research and may be incomplete.

- Data on exposures and outcomes may be limited.
Data on disease severity, details of intervention and diagnostic coding may be inconsistent

Data relating to possible confounders may be inadequate or missing
Lessons Learned

- Criteria to evaluate screening effectiveness
- Natural history of disease and screening
- Reviewed designs to evaluate:
  - Screening programs, health services and clinical interventions
- Bias in screening
- Measurement of health interventions can affect how we determine effectiveness
The End!
Course Evaluations are open

- Your feedback is important to us!
- We take these reviews into account when we revise the course each year – please help us
- You will get a course grade that is released upon completion of your course evaluation after make-up exams are completed
Dr. Deal and I would like to give thanks to --
Teaching Assistants

Filip Pirsl
Taylor Smull
Frances Wang

Jongyeon (Jay) Kim
Minghao Kou
Tong Yu
Teaching Assistants

Pablo Martinez
*Lead TA*

Jiajun Wen
Faculty Lab Instructors

Kit Carson

Patti Ephraim

Eliseo Guallar

Michael Marrone

Nicholas Reed

David Shade
Guest Lecturers

Elizabeth Platz

Priya Duggal

Michel Ibrahim
Course Coordinator

Allyn Arnold

Assistant Course Coordinator

Julie Thorne
Thank You!!!!

- To you, Epi 721 students!
- You were a wonderful class! 😊
- Consider Epi 722!!