ESTIMATING OVERDIAGNOSIS FROM TRIALS AND POPULATIONS

OVERCOMING CHALLENGES, AVOIDING MISTAKES
Goal: establish conditions for valid estimation of overdiagnosis

Excess incidence or the empirical approach
- Clinical trials – continuous-screen and stop-screen
- Population studies
- Published studies

Modelling approach
- The idea of modelling
- Prerequisites for modelling
- Published studies
Incidence pattern after screening starts:

- Incidence excesses (+) followed by corresponding deficits (-)
- Excesses: screening pulls cases from the future
- Deficits: cases screen detected no longer in prevalent pool

Note: Bump in incidence observed even if there is no overdiagnosis!
TWO APPROACHES TO ESTIMATING OVERDIAGNOSIS
SYMPTOM VERSUS CAUSE

- Excess incidence
- Empirically based
- Calculate incidence with screening minus incidence without screening

- Modeling approach
- Learn about latent disease process
- Superimpose screening and derive implications for overdiagnosis
GETTING EXCESS INCIDENCE RIGHT – CLINICAL TRIALS

I. CONTINUED SCREEN TRIAL

Hypothetical setting:
- Constant preclinical incidence
- Maximum preclinical period = 6 y
- Constant test sensitivity
- No overdiagnosis

Practice of Epidemiology

Conditions for Valid Empirical Estimates of Cancer Overdiagnosis in Randomized Trials and Population Studies

Roman Gulati, Eric J. Feuer, and Ruth Etzioni
In the Trial setting, equal numbers of individuals are randomized to a screen or control arm. Latent disease onset occurs at a constant rate. The preclinical duration follows a uniform distribution. The control arm receives no screen tests, so control arm incidence matches latent onset in this arm. The screen arm begins screening in year 1. The empirical difference between screen and control arm incidence can provide an unbiased estimate of the number of overdiagnosed cases under 2 conditions. (1) The difference is based on cumulative incidence if the screen arm stops screening and on annual incidence if the screen arm continues screening. (2) The difference is calculated after screening stabilizes plus the maximum preclinical duration.
THE PROBLEM WITH CUMULATIVE EXCESS INCIDENCE

What we know

- Screening interval
- Corresponding cases in the absence of screening
- Cases detected under screening

What we observe

- Screening interval
- Corresponding cases in the absence of screening
- Cases detected under screening

In the continued-screen setting cumulative excess incidence will be greater than zero even if NO overdiagnosis!
GETTING EXCESS INCIDENCE RIGHT – CLINICAL TRIALS

II. STOP SCREEN TRIAL

Hypothetical setting:

- Constant preclinical incidence
- Maximum preclinical period = 6 y
- Constant test sensitivity
- No overdiagnosis

Practice of Epidemiology

Conditions for Valid Empirical Estimates of Cancer Overdiagnosis in Randomized Trials and Population Studies

Roman Gulati®, Eric J. Feuer, and Ruth Etzioni
**POPULATION STUDIES**

- Background incidence generally not available – no control group
- As in clinical trials – cumulative excess incidence is persistently biased
- Annual excess incidence – wait until screening stabilizes plus max preclin duration
In the Population setting, latent disease onset occurs at a constant rate and the preclinical duration follows a uniform distribution. Initially, before screening starts, the model projects disease incidence in steady state, so that diagnosis without screening matches latent onset in the population. Annual screening begins in segments of the population at specified starting years. The empirical difference between annual incidence with and without screening provides an unbiased estimate of overdiagnosis after screening stabilizes plus the maximum preclinical duration.

**Input parameters**

- **Population size:**
  - 1,000
  - 10,000
  - 100,000

- **Annual rate of onset:**
  - 0.001

- **Range of preclinical durations:**
  - 0
  - 3

- **Episode sensitivity:**
  - 0.6

**Disease incidence**

Annual number of cases

Year
CONDITIONS FOR VALID EMPIRICAL ESTIMATES OF OVERDIAGNOSIS

**Cumulative excess incidence**
- Continued-screen trials and population settings: persistently biased
- Stop-screen trials: wait until end of screening interval plus maximum preclinical duration

**Annual (point) excess incidence**
- Continued-screen trials: unbiased at end of maximum preclinical duration
- Stop-screen trials: unbiased at end of screening interval plus max preclin duration
- Population setting: unbiased at end of screening stabilization plus max preclin duration

- In all cases: take note of denominator used in calculating overdiagnosis
Screening and Prostate-Cancer Mortality in a Randomized European Study

Fritz H. Schröder, M.D., Jonas Hugosson, M.D., Monique J. Roobol, Ph.D.,

- Cumulative excess incidence
- Continued-screen trial

<table>
<thead>
<tr>
<th>Year of publication</th>
<th>Median follow-up, years</th>
<th>Overdiagnosis among screen detections</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>9</td>
<td>58%</td>
</tr>
<tr>
<td>2012</td>
<td>11</td>
<td>55%</td>
</tr>
<tr>
<td>2014</td>
<td>13</td>
<td>49%</td>
</tr>
</tbody>
</table>
## CANADIAN NATIONAL BREAST SCREENING STUDY

### Cumulative Incidence of Invasive Cancers

<table>
<thead>
<tr>
<th>Trial arm</th>
<th>N</th>
<th>Cumulative incidence of invasive cancers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Years 1-5</td>
</tr>
<tr>
<td>Mammography+CBE</td>
<td>44,925</td>
<td>666</td>
</tr>
<tr>
<td>CBE only</td>
<td>44,910</td>
<td>524</td>
</tr>
<tr>
<td>Excess cancers in mammography arm</td>
<td>142</td>
<td>100</td>
</tr>
<tr>
<td>Excess among 484 screen detections</td>
<td>29%</td>
<td>21%</td>
</tr>
</tbody>
</table>

- Cumulative excess incidence
- Stop-screen trial

CNBSS Includes years after trial screens

Miller et al, BMJ, 2014
Most provinces started screening programs soon after trial screens ended.

### Table 3
CANBSS participants by province, years participated, and year organized provincial screening programs commenced

<table>
<thead>
<tr>
<th>Province</th>
<th>CNBSS participants</th>
<th>Years participants</th>
<th>Year provincial program commenced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Entered CNBSS</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>89835</strong></td>
<td><strong>100.0</strong></td>
<td><strong>1980–85</strong></td>
</tr>
</tbody>
</table>

More screening in mammography arm after trial screens?
PROSTATE CANCER INCIDENCE IN THE US POPULATION

Since 1986, an estimated additional 1,305,600 men were diagnosed with prostate cancer

- Cumulative excess incidence
- Background incidence imputed based on incidence in years prior to screening
BREAST CANCER INCIDENCE IN THE US POPULATION

Women aged 40 and older

0.25% increase per year based on under 40 trends

31% of detected cancers in 2008 overdiagnosed

- Point excess incidence
- Background incidence imputed based on incidence trends in women under 40
Prostate Cancer Incidence: White Males (SEER 9 Registries)
BREAST CANCER INCIDENCE IN NORWAY

15-20% overdiagnosis relative to incidence expected in absence of screening

- Cumulative excess incidence after 1st yr
- Background incidence imputed based on counties not implementing screening
**WHAT IS THE MAXIMUM PRECLINICAL DURATION FOR INVASIVE BREAST CANCER?**

*Screening Sensitivity and Sojourn Time From Breast Cancer Early Detection Clinical Trials: Mammograms and Physical Examinations*

By Yu Shen and Marvin Zelen

<table>
<thead>
<tr>
<th>Location</th>
<th>MST (Years)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIP</td>
<td>2.5</td>
<td>1.2</td>
</tr>
<tr>
<td>Sweden</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malmö</td>
<td>5.5</td>
<td>2.1</td>
</tr>
<tr>
<td>Stockholm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>2.1</td>
<td>1.3</td>
</tr>
<tr>
<td>50-59</td>
<td>2.6</td>
<td>.61</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>4.3</td>
<td>.37</td>
</tr>
<tr>
<td>Canada 1</td>
<td>1.9</td>
<td>1.2</td>
</tr>
<tr>
<td>Canada 2</td>
<td>3.1</td>
<td>.94</td>
</tr>
</tbody>
</table>

*Consensus mean 3.3 (~40 months)*

*JCO 2001*
1. Go beyond observed data to learn about underlying disease process
   - Given data on screening uptake
   - Use incidence before and after screening to learn about disease natural history

GOING BEYOND THE DATA
USING MODELING TO LEARN ABOUT OVERDIAGNOSIS
1. Go beyond observed data to learn about underlying disease process
   - Given data on screening uptake
   - Use incidence before and after screening to learn about disease natural history

   - Infer overdiagnosis based on the estimated natural history
     - Overdiagnosis occurs when other-cause death happens before the data of clinical diagnosis
PREREQUISITES FOR A USEFUL MODEL

A. Need data on disease incidence with and without screening
   - Screening trials: control group provides the counterfactual incidence
   - Population studies: may need to guesstimate a counterfactual

B. Need information on screening patterns that produced the incidence
   - Screening trials: have individual-level data on screening and mode of diagnosis
   - Population studies: typically have to reconstruct screening trends; individual-level data generally not available

C. Need a model that is identifiable (estimable) from the data
THE IDENTIFIABILITY PROBLEM
CAN THE MODEL BE LEARNED FROM THE DATA?

A) Progressive disease only

\[ S_0 \rightarrow S_p \rightarrow S_c \]

- Three parameters:
  - Risk of onset
  - Risk of progression to clinical dx
  - Screening test sensitivity

B) Mixture of progressive and indolent disease

\[ S_0 \]

\[ S_p \]

\[ S_c \]

\[ S_{p'} \]

- Indolent cases
- Four parameters:
  - Risk of onset
  - Risk of being indolent
  - If not: Risk of progression to clinical dx
  - Screening test sensitivity

*Can be learned from incidence with and without screening given screening patterns*
A SIMPLE EXPERIMENT OF IDENTIFIABILITY

In a survival analysis dataset with data censored at 5 years, the following underlying models are all consistent with the data:

- Exponential mean 40 months
- Mixture of 75% exponential with mean 18 months, 25%(effectively) infinite
- Mixture of 95% exponential with mean 26 months, 5% infinite

All will yield a mean of 40 months under an exponential model. Different models are equally consistent with the same data.

Etzioni & Gulati, JNCI 2016
BREAST CANCER NATURAL HISTORY FROM A TRIAL

A. Counterfactual incidence from a control group
B. Individual level screening histories
C. Progressive disease assumption – exponential sojourn time assumed while screening test sensitivity is estimated

A) Progressive disease only

\[ S_0 \rightarrow S_p \rightarrow S_c \]
A. Assume incidence in the absence of screening would have remained constant at pre-PSA rates
B. Aggregate screening histories retrospectively constructed from NHIS and SEER-Medicare
C. Progressive disease assumption – risk of progression to advanced or symptomatic disease depends on PSA growth rate which varies across men based on data from the PCPT trial

\[
\begin{align*}
S_0 \rightarrow S_p \rightarrow S_c
\end{align*}
\]
A FLEXIBLE PROGRESSIVE DISEASE MODEL YIELDS HETEROGENEITY IN SOJOURN TIMES

Distributions of sojourn times from a population model of prostate cancer

- **Relevant:** diagnosed within lifetime
- **Uncensored:** indolent until death

Sojourn times for relevant cancers are shorter in older men to ensure diagnosis before death

While a mixture model is not explicitly assumed, the model structure builds in heterogeneity
A. Counterfactual incidence from a control group – constant over interval analyzed

B. Individual level screening histories

C. Model allows for non-progressive disease but for identifiability **needs to assume known test sensitivity**

B) Mixture of progressive and indolent disease
IDENTIFYING IDENTIFIABILITY (OR LACK THEREOF) CAN BE HARD

A. Counterfactual incidence extrapolated from an age-period-cohort model – increasing over the screening interval

B. Aggregate, reconstructed screening histories

C. Each model has a different structure and method for estimating parameters
Overdiagnosis is complex – estimation must satisfy certain criteria to be valid

Empirical approach – excess incidence

- **Design** – stop screen or continued-screen?
- **Estimate** – cumulative or point excess incidence?
- **Timing** - has enough time elapsed?
- **Counterfactual** - Is the control group a fitting counterfactual?

Modeling approach

- **Screening patterns** – are these properly informed by available data?
- **Counterfactual** – what is the counterfactual in a population setting?
- **Identifiability** – how is the model constructed to permit identifiability?
ACKNOWLEDGMENTS

- Roman Gulati
- Lurdes Inoue
- Eric Feuer (NCI)
- CISNET support