HOW MUCH DOES SCREENING REDUCE CANCER MORTALITY?

James A. Hanley¹, Zhihui (Amy) Liu¹,², Nandini Dendukuri¹,³, Erin Strumpf¹,⁴

¹Dept. of Epidemiology, Biostatistics & Occupational Health
²Cancer Care Ontario
³Dept. of Medicine, and Technology Assessment Unit
⁴Dept. of Economics

McGill University

51ST ANNUAL ANDRÉ AISENSTADT MEMORIAL CLINICAL DAY
CANCER SCREENING – UPDATE 2014

A Symposium in Honour of Dr. André Lisbona
Jewish General Hospital

October 22, 2014
Comparison of response evaluation in small cell lung cancer using computerized tomography and chest radiography.

PMID: 8081704 [PubMed - indexed for MEDLINE]
Related citations

The diagnostic and prognostic value of renal allograft biopsy.

Parfrey PS, Kuo YL, Hanley JA, Knaack J, Xue Z, Lisbona R, Gutmann RD.
• Harms have been (well) measured; benefits have been mis-measured
• By ignoring the delay until the reductions in mortality are expressed, the prevailing interpretations of the results of cancer screening trials under-estimate the mortality reductions that would be produced by a sustained screening program
• P-value-driven RCT stopping/reporting rules exacerbate the problem
• Ways we might be able to avoid such misleading estimates
• Lung, Prostate, Colon: re-analysis of data from trials
• Breast: data from outdated trials population-screening
Why do so many trials yield a 20% ‘mortality reduction’? [Theorem]

The mortality reductions produced by a cancer screening program

A way ahead? (impact of N-round program: $\sum_{i=1}^{N} \text{impact of round}_i$)

Illustrations: cancer of the prostate, breast, colon

Comments: cancer of the breast
20% MORTALITY REDUCTION

A UNIVERSAL CONSTANT IN CANCER SCREENING TRIALS?
For many RCTs, single rate (hazard) ratio or risk difference is OK

- A single (overall) Rate Reduction (i.e., single Rate Ratio), based on all events that have occurred (regardless of when) up to end of available follow-up time on each subject

- ‘Regardless of when’ implies proportional hazards, i.e., reduction is immediate & sustained (if need be, by continuing to take medications)

- **Numbers of events matter, but not their timing:** Q: how to have sufficient events for desired precision?  
  more persons, less time? ↔ more time, fewer persons?

- As amount of person time (number of events) increases, updated single Rate Reduction traces out a random walk
Reductions in ‘event rates’ as follow-up time unfolds

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

Cumulative Number of Events

(P = 0.05)

50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
Examples of ‘prevention’ / ‘early detection’ studies
HIV: if ‘intervention’ ineffective

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

---

---

---
HIV: Adult circumcision

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

Rate Ratio

--- Time

(Cumulative) Number of Events

- P = 0.05
- 50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
Paralytic or non-paralytic poliomyelitis: Salk Vaccine

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

2) Paralytic/non-paralytic poliomyelitis
[Salk Vaccine]
HPV_{6,11,16,18} infection: Quadrivalent HPV Vaccine

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

1) Human Papilloma Virus (HPV) infection
   [Quadrivalent HPV vaccine]

2) Paralytic/non-paralytic poliomyelitis
   [Salk Vaccine]

P = 0.05
Death from ruptured abdominal aortic aneurysm: Ultrasound screening

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

1) Human Papilloma Virus (HPV) infection
   [Quadrivalent HPV vaccine]
2) Paralytic/non-paralytic poliomyelitis
   [Salk Vaccine]
3) Death from ruptured abdominal aortic aneurysm
   [Ultrasound screening]

P = 0.05

(Cumulative) Number of Events

0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Cancer Screening Trial - theoretical

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

P = 0.05

Rate Ratio

---→ Time

(Cumulative) Number of Events

0%
2%
14%
35%
50%
50%
50%
7)

→ Time 0
0.5
1 [ref.]

[0%]
[2%]
[14%]
[35%]
[50%]
[50%]
[50%]
3 actual cancer screening trials

Percentage Reduction in Average Event Rate, if data are analyzed after indicated no. of events

1) Human Papilloma Virus (HPV) infection
   [Quadrivalent HPV vaccine]

2) Paralytic/non-paralytic poliomyelitis
   [Salk Vaccine]

3) Death from ruptured abdominal aortic aneurysm
   [Ultrasound screening]

4) PROSTATE
5) COLON
6) LUNG
(a) Age-specific numbers of prostate cancer deaths in a steady state population with a given age-structure, if screening had not been available, and if screening had been available from ages 50 to 70.
or (b) the Rate Ratio (or %Reduction) Function ...

(a) Age-specific numbers of prostate cancer deaths in a steady state population with a given age-structure, if screening had not been available, and if screening had been available from ages 50 to 70

(b) The corresponding age-specific prostate cancer mortality rate ratios
‘% Reduction function’ (bathtub shape)

- The asymptote is the ultimate estimand

- It is determined by ...
  - number and spacing of rounds, and
  - the contribution of each round of screening

- From published trials, can one ..
  - estimate the % Reduction function ?
  - estimate contribution of each round ?
    (?? function shape if different schedule or if a program)
PROSTATE CANCER
Screening & Prostate-Ca Mortality in Randomized European Study ’92-’08 (“ERSPC” nejm2009.04)

As of December 31, 2006, with an average follow-up time of 8.8 years, there were 214 prostate-cancer deaths in the screening group and 326 in the control group. (...) The adjusted rate ratio for death from prostate cancer in the screening group was 0.80 (95% CI, 0.65 to 0.98; P=0.04).

“PSA-based screening reduced the rate of death from prostate cancer by 20%.”
RE-ANALYSIS OF ERSPC DATA
using
year-specific prostate cancer mortality ratios
(A) Overall vs. (B) Year-specific mortality ratios

(B) Prostate Cancer Mortality Rate Ratio (S / C)

Cumulative Prostate Cancer Mortality

European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up


Summary
Background The European Randomised study of Screening for Prostate Cancer (ERSPC) has shown significant reductions in prostate cancer mortality after 9 years and 11 years of follow-up, but screening is controversial because of adverse events such as overdiagnosis. We provide updated results of mortality from prostate cancer with follow-up to 2010, with analyses truncated at 9, 11, and 13 years.

Methods ERSPC is a multicentre, randomised trial with a predefined centralised database, analysis plan, and core age group (55–69 years), which assesses prostate-specific antigen (PSA) testing in eight European countries. Eligible men aged 50–74 years were identified from population registries and randomly assigned by computer generated random numbers to screening or no intervention (control). Investigators were masked to group allocation. The primary outcome was prostate cancer mortality in the core age group. Analysis was by intention to treat. We did a secondary analysis that corrected for selection bias due to non-participation. Only incidence and no mortality data at 9 years’ follow-up are reported for the French centres. This study is registered with Current Controlled Trials, number ISRCTN49127736.

Findings With data truncated at 13 years of follow-up, 7408 prostate cancer cases were diagnosed in the intervention group and 6107 cases in the control group. The rate ratio of prostate cancer incidence between the intervention and control groups was 1.91 (95% CI 1.83–1.99) after 9 years (1.64 [1.58–1.69] including France), 1.66 (1.60–1.73) after 11 years, and 1.57 (1.51–1.62) after 13 years. The rate ratio of prostate cancer mortality was 0.85 (0.70–1.03) after 9 years, 0.78 (0.66–0.91) after 11 years, and 0.79 (0.69–0.91) at 13 years. The absolute risk reduction of death from prostate cancer at 13 years was 0.11 per 1000 person-years or 1.28 per 1000 men randomised, which is equivalent to one prostate cancer death averted per 781 (95% CI 490–1929) men invited for screening or one per 27 (17–66) additional prostate cancer detected. After adjustment for non-participation, the rate ratio of prostate cancer mortality in men screened was 0.73 (95% CI 0.61–0.88).

Interpretation In this update the ERSPC confirms a substantial reduction in prostate cancer mortality attributable to testing of PSA, with a substantially increased absolute effect at 13 years compared with findings after 9 and 11 years. Despite our findings, further quantification of harms and their reduction are still considered a prerequisite for the introduction of populated-based screening.
Role of the funding source
The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. Access to data was limited to the independent data centre led by SMM. None of the investigators had access to outcome data outside the planned official reports of the data centre. FHS produced the primary version and was responsible for submitting the report.

Results
In the core group of men aged 55–69 years, excluding France, 162 388 were randomly assigned, of whom 145 died between randomisation and screening. With data truncated at 13 years of follow-up, 7408 prostate cancer deaths were reported. The intervention group had a lower rate of prostate cancer deaths compared to the control group during all time periods.

Table 3: Prostate cancer mortality in the intervention and control groups during three time periods truncated (all centres, core age group, France excluded except for years 1–9)

<table>
<thead>
<tr>
<th>Years</th>
<th>Person-years</th>
<th>Rate per 1000 person-years</th>
<th>Prostate cancer deaths</th>
<th>Person-years</th>
<th>Rate per 1000 person-years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years 1–9</td>
<td>193 614</td>
<td>0·31</td>
<td>590</td>
<td>278 751</td>
<td>0·37</td>
</tr>
<tr>
<td>Years 1–9 including France</td>
<td>7902 835</td>
<td>9·46</td>
<td>353</td>
<td>5726 984</td>
<td>5·81</td>
</tr>
<tr>
<td>Years 1–11</td>
<td>265 732</td>
<td>0·35</td>
<td>133</td>
<td>415 896</td>
<td>0·46</td>
</tr>
<tr>
<td>Years 1–13</td>
<td>355 825</td>
<td>0·43</td>
<td>018</td>
<td>545 1 011</td>
<td>0·54</td>
</tr>
</tbody>
</table>

*Adjusted by centre and for the randomisation ratio 1:1·5 intervention group versus control group in Finland.

Figure 2: Nelson–Aalen estimates of cumulative prostate cancer mortality (all centres, excluding France)
Figure 3: Nelson-Aalen estimates of cumulative prostate cancer in both groups by 4-year periods (all centres, excluding France)
BREAST CANCER

EVERY TRIAL & META-ANALYSIS:

and (nejm2010) REPORT on NORWAY NATIONAL SCREENING PROGRAM:

REDUCTION UNDER-ESTIMATED

### Observed breast cancer mortality deficits in 5 Mammography Trials

<table>
<thead>
<tr>
<th>Study</th>
<th>Ages at entry</th>
<th>Ratio of No. in Experimental arm : No. in Control arm</th>
<th>Participation Rate in screens that were part of trial</th>
<th>Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H.I.P.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-69</td>
<td>1 : 1</td>
<td>65%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 : 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Malmo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55-70</td>
<td>1 : 1</td>
<td>70%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 : 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gothenburg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39-59</td>
<td>0.7 : 1</td>
<td>82%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2 Counties</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-74</td>
<td>1.4 : 1</td>
<td>89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stockholm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-64</td>
<td>2 : 1</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Year-specific data**: trials used by Task Force.
- **20 years of screening**, 50–69, would be followed by 20 years (55–74) in which the breast cancer mortality reduction in these years would be $\geq 40\%$, with smaller deficits in other years.
- Fewer than 200 women would need to participate in such a program in order to avert a breast cancer death in the age range 50-80.

**Corresponding Task Force estimates:**

- Mortality reduction: 21%;
- Number of women: 720.
COLON CANCER
Long-Term Mortality after Screening for Colorectal Cancer

Aasma Shaukat, M.D., M.P.H., Steven J. Mongin, M.S., Mindy S. Geisser, M.S.,
Frank A. Lederle, M.D., John H. Bond, M.D., Jack S. Mandel, Ph.D., M.P.H.,
and Timothy R. Church, Ph.D.

ABSTRACT

BACKGROUND
In randomized trials, fecal occult-blood testing reduces mortality from colorectal cancer. However, the duration of the benefit is unknown, as are the effects specific to age and sex.

METHODS
In the Minnesota Colon Cancer Control Study, 46,551 participants, 50 to 80 years of age, were randomly assigned to usual care (control) or to annual or biennial screening with fecal occult-blood testing. Screening was performed from 1976 through 1982 and from 1986 through 1992. We used the National Death Index to obtain updated information on the vital status of participants and to determine causes of death through 2008.
RESULTS
Through 30 years of follow-up, 33,020 participants (70.9%) died. A total of 732 deaths were attributed to colorectal cancer: 200 of the 11,072 deaths (1.8%) in the annual-screening group, 237 of the 11,004 deaths (2.2%) in the biennial-screening group, and 295 of the 10,944 deaths (2.7%) in the control group. Screening reduced colorectal-cancer mortality (relative risk with annual screening, 0.68; confidence interval [CI], 0.56 to 0.82; relative risk with biennial screening, 0.78; 0.65 to 0.93) through 30 years of follow-up. No reduction was observed in all-cause mortality (relative risk with annual screening, 1.00; 95% CI, 0.99 to 1.01; relative risk with biennial screening, 0.99; 95% CI, 0.98 to 1.01). The reduction in colorectal-cancer mortality was larger for men than for women in the biennial-screening group (P=0.04 for interaction).

CONCLUSIONS
The effect of screening with fecal occult-blood testing on colorectal-cancer mortality persists after 30 years but does not influence all-cause mortality. The sustained reduction in colorectal-cancer mortality supports the effect of polypectomy. (Funded by the Veterans Affairs Merit Review Award Program and others.)
adjusted relative-risk estimates for death from colorectal cancer for the annual-screening and biennial-screening groups were 0.65 (95% CI, 0.52 to 0.80) and 0.76 (95% CI, 0.61 to 0.95), respectively.

Annual or biennial screening with fecal occult-blood testing had no apparent effect on all-cause mortality. The relative risk of death from any cause was 1.00 (95% CI, 0.99 to 1.01) with annual screening, 0.99 (95% CI, 0.98 to 1.01) with biennial screening, and 1.00 (95% CI, 0.98 to 1.01) with annual and biennial screening combined (Fig. 2 and Table 1). No effect was seen on deaths from causes other than colorectal cancer; the relative risk of death from causes unrelated to colorectal cancer was 1.00 (95% CI, 0.99 to 1.02) with annual screening, 1.00 (95% CI, 0.98 to 1.01) with biennial screening, and 1.00 (95% CI, 0.99 to 1.01) with annual and biennial screening combined (Fig. S5 in the Supplementary Appendix).

The causes of death are provided in Table S1 in the Supplementary Appendix.

SUBGROUP ANALYSES

Figure 3 shows the numbers of participants who underwent randomization, the numbers of those who died from colorectal cancer, and the relative risks for the subgroups of age and sex, according to each study group and the combined screening groups. Graphs of cumulative colorectal-cancer mortality and corresponding relative risks for the subgroups are shown in Figures S6 and S7 in the Supplementary Appendix. The reduction in colorectal-cancer mortality was larger for men than for women in both screening groups and in the two groups combined; the relative risk of death from colorectal cancer was 0.61 (95% CI, 0.47 to 0.80) for men vs. 0.75 (95% CI, 0.57 to 0.97) for women in the annual-screening group, 0.63 (95% CI, 0.48 to 0.82) vs. 0.92 (95% CI, 0.72 to 1.18) in the biennial-screening group, and 0.62 (95% CI, 0.50 to 0.78) vs. 0.83 (95% CI, 0.67 to 1.04) in the combined screening groups. The interaction between sex and screening, as measured by the ratio of the relative risk for men to that for women, was significant in the biennial-screening group (P = 0.04 for interaction) but not in the annual-screening group or the two groups combined (P = 0.30 and P = 0.06, respectively, for interaction).
The fairness of the draft lottery was immediately debated. Critics contended that the process was not truly random. A *New York Times* article quoted a White House source as saying “discussions that the lottery was not random are purely speculative.” In that same *New York Times* article, Senator Edward Kennedy was quoted as asking the National Sciences the “apparent lack of randomness” in the selection.

The Data

The data is publicly available on the internet. One source is the Data and Story Library. The draft lottery data is located at the following URL:

http://lib.stat.cmu.edu/DASL/Datafiles/DraftLottery.html

If you have not imported data into R from external sources, you might want to first work through the activity Importing Data in R.

One technique, as explained in Importing Data in R, suggests copying the data into a plain text file. Open a simple text editor (e.g., Notepad on Windows or Textedit on the Mac). Copy and paste the lottery data from the above URL, including headers (but not the descriptive information above the headers), and save the file as

Figure 1. Rep. Alexander Pirnie, R-NY, draws the first capsule in the lottery drawing held on Dec. 1, 1969. The capsule contained the date, Sept. 14.
In the activity Boxplots in R we learned how to use R's `boxplot` command to produce a boxplot of a data set. To examine the "fairness" of the Selective Service's draft lottery, we will produce "side-by-side" boxplots for each month of the year. That is, we will produce 12 boxplots, one for each month of the year, each containing an analysis of the associated draft numbers for that month. The following command will produce these "side-by-side" boxplots shown in Figure 4.

```r
> boxplot(Draft_No. ~ Month, data=lottery)
```

Figure 4. Side-by-side boxplots of draft numbers for each month.

Because the data in `Month` is categorical (you can see this by typing `lottery$Month`), the model formula `Draft_No. ~ Month` causes the boxplot command to group the numerical data in `Draft_No.` according to the categories in `Month`. Therefore, the command `boxplot(Draft_No. ~ Month, data=lottery)` creates 12 boxplots, one for each month. For example, the boxplot for April (see `Apr` in Figure 4) contains an analysis for only those draft numbers that were assigned to birth-dates in April. Similar comments are in order for the remaining months.

Unfortunately, the months are sorted in alphabetical order (the default behavior). It would be more appropriate if they were sorted in chronological order, January first, February second, etc. One solution would be to boxplot the draft numbers versus the month number.

```r
> boxplot(Draft_No. ~ Mo.Number, data=lottery)
```

This command produces the side-by-side boxplots shown in Figure 5.

---

*Figure 4. Side-by-side boxplots of draft numbers for each month.*
Figure 2. A scatterplot of Draft_No. versus Day_of_year.
Interpretation of Results

The image in Figure 6 is perfect. The months are now sorted in chronological order. But now, what does the image of side-by-side boxplots tell us?

Remember, the heavy horizontal bar in each box is the median of the data set. The median draft number for the month of December is very disconcerting. Remember, the lower the draft number, the more likely you would be inducted to serve in Vietnam. Why does the month of December have a median that is significantly lower than most of the other months. It seems that the men with birthdays in December are being unfairly selected. Indeed, with the exception of October, the last remaining months of the year all have medians that are significantly lower than the medians of the previous months. Something strange is going on!

One story offers a hint of an explanation. It seems that the capsules containing birthdays for January were placed in a shoe-box, thoroughly mixed, then poured into the glass container shown in Figure 1. Then the same procedure was followed for the capsules containing birthdays in February, stirring them thoroughly in a shoe-box, then pouring them into the glass container. This same procedure was followed for the remaining months.

December was the last month processed, or so the story goes.

However, this is quite disturbing. If capsules were selected from the top of the glass container, they were more likely to be a December birthday. According to the story, the person making the draws did not always reach deep into the pile of capsules. This may be one explanation for why so many December birthdays were selected early in the process and assigned low draft numbers (which correlates to a higher chance of being drafted).

This story may be an oversimplification. Readers are encouraged to explore the reasons for why this process.
Time-split versus time-lumped Rate Ratios

Rate Ratio

1-year bins

Rate Ratio

2-year bins

Rate Ratio

3-year bins

Rate Ratio

5-year bins

Rate Ratio

10-year bins

Rate Ratio

10-year bins

Rate Ratio

30-year bins
Dear Editor

- Shaukat et al. report reductions of 32% and 22% in colon cancer mortality in those offered 11 annual and 6 biennial FOB screens, respectively. These reductions were achieved despite a 4-year hiatus in screening, and averaging over all 30-years of follow-up.

- What would the reductions have been without such an interruption? To answer this, we extracted the yearly numbers of deaths from the published Figure 1, and instead calculated yearly mortality reductions. Because of the unusual schedule, the resulting reduction curve has a ‘W’ shape, showing the lagged responses to the two phases of screening: after a delay of some years, mortality reductions reached a nadir of around 40% before reverting to what they would be in the absence of screening; this pattern is repeated when screening is resumed.

- Without the (funding related) hiatus, the reductions would have been around 40% for each year affected, which is substantially larger than those estimated.
Yearly reductions in colon cancer mortality in two screening arms. Each dot is based on number of deaths in a three year moving window; smooth curves were fitted through them. Because the hiatus was in calendar-time rather than follow-up time, and entries were staggered, the timing of the screens (each denoted by an ‘S’) is only approximate.
From Trial to Program

STATISTICAL MODEL
Convolution of reductions produced by individual rounds

If one round of screening reduces mortality in each of 5 future years, then in a trial, 3 rounds of screening -- S1, S2 and S3 -- would produce 3 ‘waves’ of mortality reductions (‘1’, ‘2’, ‘3’), each 5 years wide, over 7 years (Y3-Y9).

In such a trial, the maximal reduction (35%, year 6) would be smaller than the sustained (46%) reductions produced by a 20-year screening program.

The average reduction, computed over 13 years of follow-up in such a trial would be an even more serious underestimate of the impact of a 20-year program.
Figure 5–4: Panel A: Empirical and fitted mortality reductions based on the yearly numbers of colorectal cancer deaths in the two screening arms of the Minnesota Colorectal Cancer Screening Study, with the 4-year hiatus. The size of each dot is proportional to the information contribution of the empirical year-specific mortality ratio. Because the hiatus was in calendar-time rather than follow-up time, and entries were staggered, the timing of the screens, each denoted by an S, is only approximate.

Panel B: Projection of yearly mortality reductions in colorectal cancer that would be generated by 15 years of uninterrupted annual and biennial fecal occult blood screening. The grey area represents time-specific 95% confidence bands under the biennial screening regimen.
LUNG CANCER
Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening

The National Lung Screening Trial Research Team*

ABSTRACT

BACKGROUND
The aggressive and heterogeneous nature of lung cancer has thwarted efforts to reduce mortality from this cancer through the use of screening. The advent of low-dose helical computed tomography (CT) altered the landscape of lung-cancer screening, with studies indicating that low-dose CT detects many tumors at early stages. The National Lung Screening Trial (NLST) was conducted to determine whether screening with low-dose CT could reduce mortality from lung cancer.

METHODS
From August 2002 through April 2004, we enrolled 53,454 persons at high risk for lung cancer at 33 U.S. medical centers. Participants were randomly assigned to undergo three annual screenings with either low-dose CT (26,722 participants) or single-view posteroanterior chest radiography (26,732). Data were collected on cases of lung cancer and deaths from lung cancer that occurred through December 31, 2009.

RESULTS
The rate of adherence to screening was more than 90%. The rate of positive screen-
NLST
Age at entry: 55–74
CT: X-ray allocation = 1 : 1
Compliance = 94%

Figure 6–1: NLST yearly numbers of lung cancer deaths, extracted from published NEJM report.
Figure 6–2: NLST yearly numbers of lung cancer deaths, with relatively large hypothetical reductions in years 7-10.
Figure 6–3: NLST yearly numbers of lung cancer deaths, with relatively small hypothetical reductions in years 7-10.
Figure 6–4: NLST number at risk for the two arms, along with lung cancer deaths, using the individual-level data provided by the NCI.

Checking the yearly numbers that we extracted in Table 6–1(a) against those calculated from individual-level data in Table 6–1(b) was one of our first tasks, by including lung cancer deaths before the cutoff date only. They were almost identical, only differing by a few deaths. Next we included lung cancer deaths also after the cutoff date.
Table 6–1: Yearly numbers of lung cancer deaths in the NLST. Part (a) was based on our extraction from the NEJM report, (b) and (c) are based on the individual-level NLST data; in (b) only deaths that occurred before the cut-off (i.e. January 15th, 2009) were included, and in (c) all deaths occurred before and after the cutoff date were included.

(a) Year-specific data extracted from figure in NEJM report

<table>
<thead>
<tr>
<th>Follow-up Year:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screens</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X-ray Arm:</td>
<td>37</td>
<td>68</td>
<td>82</td>
<td>95</td>
<td>84</td>
<td>73</td>
<td>4</td>
<td>442</td>
</tr>
<tr>
<td>CT Arm:</td>
<td>31</td>
<td>57</td>
<td>67</td>
<td>84</td>
<td>72</td>
<td>42</td>
<td>3</td>
<td>354</td>
</tr>
<tr>
<td>Reduction:</td>
<td>16%</td>
<td>16%</td>
<td>18%</td>
<td>12%</td>
<td>14%</td>
<td>42%</td>
<td>25%</td>
<td>20%</td>
</tr>
</tbody>
</table>

(b) Year-specific data including deaths before the cutoff only

| X-ray Arm:      | 38| 70| 83| 91| 88| 74| 4 | 448   |
| CT Arm:         | 31| 57| 67| 84| 72| 45| 3 | 359   |
| Reduction:      | 18%| 19%| 19%| 8%| 18%| 39%| 25%| 20%   |

(c) Year-specific data including deaths before and after the cutoff

| X-ray Arm:      | 38| 70| 83| 91| 89| 116| 65| 552   |
| CT Arm:         | 31| 57| 67| 84| 73| 85| 70 | 467   |
| Reduction:      | 18%| 19%| 19%| 8%| 18%| 27%|-8%| 15%   |
Figure 6–5: NLST yearly numbers of lung cancer deaths, correspond 6–1(c).
be around 30%, which doubles the 15% reduction achieved with three rounds of screening in the trial. One could easily study whether a younger or older age group would benefit more from early detection, by choosing data on those aged, say, 65 years or younger at randomization. The fitted curve and the corresponding projection based on 10 rounds of annual screening are shown in Figure 6–6. Our choice of the age group is rather arbitrary, but this serves as an illustration for other subgroup analyses, such as splitting by gender, ethnicity group, medical history and so on.

Figure 6–6: Fitted reduction curve (dotted, black) based on the NLST data for persons aged below 65 at onset of screening and projected curve based on 10 rounds of annual screenings.
By ignoring the delay until the reductions in mortality are expressed, the prevailing interpretations of the results of cancer screening trials **under-estimate** the mortality reductions that **would be produced by a sustained screening program**.

P-value-driven RCT stopping/reporting rules exacerbate the problem.

We *might* be able to avoid such misleading estimates if we . . .

(i) distinguish a trial from a program
(ii) run trials with sufficient rounds of screening and sufficient follow-up
(iii) spend major portion of career waiting to measure real reductions
(iv) analyze the data using time-specificity / non-proportional hazards
(v) focus on parameters describing **impact of 1 round of screening**
(vi) mammography: use data from population-screening, not old trials
FUNDING, CO-ORDINATES, DOWNLOADS

Natural Sciences and Engineering Research Council of Canada

Le Fonds québécois de la recherche sur la nature et les technologies

Canadian Institutes of Health Research (2011-2014)

James.Hanley@McGill.CA
Zhihui.Liu@Mail.McGill.CA

www.med.mcgill.ca/epidemiology/hanley

→ reprints / talks

McGill Biostatistics
Biostatistique
Why do statisticians commonly limit their inquiries to Averages?

F. Galton, Natural Inheritance, 1889.

“It is difficult to understand why statisticians commonly limit their inquiries to Averages, and do not revel in more comprehensive views.

Their souls seem as dull to the charm of variety as that of the native of one of our flat English counties, whose retrospect of Switzerland was that, if its mountains could be thrown into its lakes, two nuisances would be got rid of at once.”
Timing of cholesterol reductions produced by statins

3 dogs at 20 mg/kg/day; 3 at 50 mg/kg/day

3 monkeys at 50 mg/kg/day
Timing of cholesterol reductions produced by statins

Humans
The loneliness of the long-distance trialist

Timing of Screening Effects
(as seen in cumulative cause-specific mortality curves)

Abdominal Aortic Aneurysms
(One-off Screening, MASS)

Prostate Cancer
(q 4y, ERSPC)

Cumulative Cause-Specific Mortality

Follow-Up Year

Supp Fig. A