Behavior within a Clinical Trial and Implications for Mammography Guidelines

Amanda E. Kowalski
Gail Wilensky Professor of Applied Economics and Public Policy
Department of Economics, University of Michigan

February 2022
“The decision to start screening mammography in women prior to age 50 years should be an individual one. Women who place a higher value on the potential benefit than the potential harms may choose to begin biennial screening between the ages of 40 and 49 years”
Do current guidelines target mammograms to women most likely to benefit from them?
“The most important harm is the diagnosis and treatment of noninvasive breast cancer that would otherwise not have become a threat to a woman's health, or even apparent, during her lifetime (that is, overdiagnosis and overtreatment).” (Siu, 2016).
Overdiagnosis

Clinical trials provide the best evidence

I Examine Behavior within a Clinical Trial

• Canadian National Breast Screening Study
  – 89,835 patients enrolled
  – Intervention: access to mammograms for 4 to 5 years during active study period
  – Recorded mammograms, even in control group
  – Examined breast cancer and mortality outcomes from 1980 to 2005 (at least 20 years for all participants) through cancer registry and death records (no attrition)
  – Collected risk factors and demographic data
I Examine Behavior within a Clinical Trial

1. Heterogeneous selection: are women who are more likely to receive mammograms different from other women?

2. Treatment effect heterogeneity: are women who are more likely to receive mammograms more likely to experience better or worse health outcomes because of them?
I Examine Behavior within a Clinical Trial

• Clinical trial literature says little about mammography behavior
  – See Nelson (2016)

• Natural experiment literature examines mammography behavior
  – Many papers focus on mammography as an outcome but do not examine selection and treatment effect heterogeneity
    • Kelaher and Stellman (2000); Habermann, Virnig, Riley, Baxter (2007); Kadiyala and Strumpf (2011, 2016); Finkelstein, Taubman, Wright, Bernstein, Gurber, Newhouse, Allen, Baicker, Oregon Health Study Group (2012); Kolstad and Kowalski (2012); Bitler and Carpenter (2016, 2019); Fedewa, Goodman, Flanders, Han, Smith, Ward, Doubeni, Sauer, Jemal (2015); Mehta, Polsky, Zhu, Lewis, Kolstad, Loewenstein, Volpp (2015); Ong and Mandl (2015); Lu and Slusky (2016); Zanella and Banerjee (2016); Cooper, Kou, Dor, Koroukian, Schluchter (2017); Jacobson and Kadiyala (2017); Buchmueller and Goldzahl (2018); Myerson, Tucker-Seeley, Goldman, Lakdawalla (2020)
  – Two papers corroborate the selection heterogeneity that I find but do not examine treatment effect heterogeneity
    • Kim and Lee (2017); Einav, Finkelstein, Oostrom, Ostriker, Williams (2020)
I Examine Behavior within a Clinical Trial

- I build on LATE and MTE literatures from economics
  - Bjorklund and Moffitt (1987)
  - Imbens and Angrist (1994)
  - Vytlacil (2002)
  - Brinch, Mogstad, Wiswall (2015)
  - Cornelissen, Dustmann, Raute, Schoenberg (2018)
I Examine Behavior within a Clinical Trial


*Divided into two papers immediately below.*

“Reconciling Seemingly Contradictory Results from the Oregon Health Insurance Experiment and the Massachusetts Health Reform.”

*Accepted, Review of Economics and Statistics.*

“Behavior within a Clinical Trial and Implications for Mammography Guidelines.”

*Accepted, Review of Economic Studies.*

“How to Examine External Validity Within an Experiment.”

*Accepted, Journal of Economics and Management Strategy.*


*Extending with Jann Spiess.*
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model of Behavior within a Clinical Trial
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be overdiagnosed by them

• Robustness
• Implications for Mammography Guidelines
Fraction treated $p$
\[ Z = 0 \]

\[ D = 1 \]

\[ D = 0 \]

\[ p_c = 0.19 \]

Fraction treated \( p \)

0.00 \( \rightarrow \) \( p_c = 0.19 \) \( \rightarrow \) 1.00
\[ Z = 0 \]

\[ D = 1 \]

\[ D = 0 \]

\[ D = 0.00 \]

\[ 1.00 \]

\[ p_c = 0.19 \]

Always Takers

Fraction treated \( p \)
\[ Z = 1 \]

\[ Z = 0 \]

\[ D = 1 \]

\[ D = 0 \]

Fraction treated \( p \)

\[ 0.00 \quad p_c = 0.19 \quad 1.00 \]

\[ p_1 = 0.95 \]

Always Takers
\[ Z = 0 \]
\[ D = 0 \]
\[ Z = 1 \]
\[ D = 1 \]

\[ p_c = 0.19 \]
\[ p_t = 0.95 \]

Always Takers

Never Takers

Fraction treated \( p \)
\( Z = 1 \)  

\( D = 1 \)  

\( D = 0 \)  

\( Z = 0 \)  

\( D = 1 \)  

\( D = 0 \)  

\( \frac{17}{43} \)  

\( p_c = 0.19 \)  

\( p_I = 0.95 \)  

Always Takers  

Compliers  

Never Takers  

Fraction treated \( p \)
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model of Behavior within a Clinical Trial
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be overdiagnosed by them

• Robustness
• Implications for Mammography Guidelines
Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

- **treated**
- **untreated**

For Always Takers ($p_C = 0.19$):

\[
\begin{align*}
\text{Breast Cancer Incidence} &= 571 \\
(0.95 \times 453 - 0.19 \times 571) \\
(0.95 - 0.19)
\end{align*}
\]

For Compliers:

\[
\frac{(1-0.19) \times 385 - (1-0.95) \times 667}{(0.95 - 0.19)}
\]

For Never Takers ($p_I = 0.95$):

\[
\begin{align*}
\text{Breast Cancer Incidence} &= 667 \\
(0.95 \times 453 - 0.19 \times 571) \\
(0.95 - 0.19)
\end{align*}
\]

Fraction treated $p$:

- $Z = 0$ (D=0)
- $Z = 1$ (D=1)
Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

- **treated**
- **untreated**

$p_C = 0.19$

$p_I = 0.95$

Never Takers

Fraction treated $p$

Always Takers
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model of Behavior within a Clinical Trial
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be overdiagnosed by them

• Robustness
• Implications for Mammography Guidelines
untreated outcome test rejects selection homogeneity:

$$366 - 667 = -301 \neq 0$$

$$[0.003]$$

LATE = 58

(38)
All-Cause Mortality

untreated outcome test rejects selection homogeneity:

\[ 428 - 990 = -562 \neq 0 \]

[0.000]

LATE = -13 (39)
Women more likely to receive mammograms have higher socioeconomic status

<table>
<thead>
<tr>
<th>Baseline Socioeconomic Status</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>In work force</td>
<td>0.65 0.64 0.65</td>
</tr>
<tr>
<td>Age at first birth</td>
<td>24.28 23.98 23.57</td>
</tr>
<tr>
<td>No live birth</td>
<td>0.16 0.15 0.13</td>
</tr>
<tr>
<td>Married</td>
<td>0.80 0.81 0.75</td>
</tr>
<tr>
<td>Husband in work force if alive</td>
<td>0.81 0.81 0.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Baseline Health Behavior</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Smoker</td>
<td>0.78 0.75 0.63</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>23.87 24.42 24.48</td>
</tr>
<tr>
<td>Used oral contraception</td>
<td>0.74 0.71 0.67</td>
</tr>
<tr>
<td>Used estrogen</td>
<td>0.13 0.13 0.15</td>
</tr>
<tr>
<td>Any mammograms prior to enrollment</td>
<td>0.23 0.13 0.13</td>
</tr>
<tr>
<td>Practiced breast self-examination</td>
<td>0.47 0.44 0.38</td>
</tr>
</tbody>
</table>
Natural Experiments Corroborate Selection Heterogeneity

• Kim and Lee (2017)
• Einav, Finkelstein, Oostrom, Ostriker, Williams (2020)
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model of Behavior within a Clinical Trial
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be overdiagnosed by them

• Robustness

• Implications for Mammography Guidelines
untreated outcome test rejects selection homogeneity:

$$366 - 667 = -301 \neq 0$$

$$[0.003]$$

LATE = 58

(38)
untreated outcome test rejects selection homogeneity:

\[ 366 - 667 = -301 \neq 0 \]

\[ [0.003] \]

\[ \text{LATE} = 58 \quad (38) \]
untreated outcome test rejects selection homogeneity:

$$366 - 667 = -301 \neq 0$$

[0.003]

Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

always taker average treatment effect
lower bound = 206
(59)

upper bound

LATE = 58
(38)
test rejects treatment effect homogeneity:
\[-301 \times (206 - 58) = -44,311 < 0\]
\ bro\ {0.023}\}

untreated outcome test
rejects selection homogeneity:
\[366 - 667 = -301 \neq 0\]
\ bro\ [0.003]}

always taker
average

treatment
effect

lower bound
= 206
(59)
Women of higher socioeconomic status are exposed to increased “observational intensity” such that “they are likely to be screened more often and by means of such tests...that can detect smaller abnormalities, undergo more follow-up testing, and undergo more biopsies, and they may be served by health systems that have a lower threshold for labeling results as abnormal.”

- Welch and Fisher (2017)
Breast Cancer Characteristics Corroborate Treatment Effect Heterogeneity

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
<th>Difference in Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Always Takers</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Treated Compliers</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Tumor Size Among Breast Cancers (in mm)</td>
<td>73</td>
<td>75</td>
</tr>
<tr>
<td>Share of Invasive Breast Cancer Among Breast Cancers (%)</td>
<td>(9)</td>
<td>(7)</td>
</tr>
</tbody>
</table>
### Procedures Corroborate Treatment Effect  
#### Heterogeneity

<table>
<thead>
<tr>
<th></th>
<th>Always Takers</th>
<th>Treated Compliers</th>
<th>Difference in Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor Size Among Breast Cancers (in mm)</td>
<td>13 (1)</td>
<td>18 (2)</td>
<td>-5 (1) - (2)</td>
</tr>
<tr>
<td>Share of Invasive Breast Cancer Among Breast Cancers (%)</td>
<td>73 (9)</td>
<td>75 (7)</td>
<td>-2 (9) - (7)</td>
</tr>
<tr>
<td>Share of Mastectomy Among Breast Cancers with Mastectomy or Lumpectomy (%)</td>
<td>45 (9)</td>
<td>23 (7)</td>
<td>22 (9) - (7)</td>
</tr>
</tbody>
</table>
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model of Behavior within a Clinical Trial
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be overdiagnosed by them

• Robustness

• Implications for Mammography Guidelines
Results Are Robust Along Many Dimensions

• Alternative outcome
  – All-cause mortality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
Suggestive Evidence for All-Cause Mortality

![Graph showing suggestive evidence for all-cause mortality with statistical tests and calculations.](image)
Behavior within a Clinical Trial and Implications for Mammography Guidelines

• Model of Behavior within a Clinical Trial
  – First Stage: Mammography
  – Second Stage: Breast Cancer Incidence

• Results
  1. Selection Heterogeneity
     • Women more likely to receive mammograms are healthier
  2. Treatment Effect Heterogeneity
     • Women more likely to receive mammograms are more likely to be overdiagnosed by them

• Robustness

• Implications for Mammography Guidelines
Implications for Mammography Guidelines

U.S. Preventive Services Task Force (USPSTF) 2016 Guidelines for Women in 40’s:

“The USPSTF recommends selectively offering or providing this service to individual patients based on professional judgment and patient preferences”
Results Are Robust Along Many Dimensions

• Alternative outcome
  – All-cause mortality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
## Main Specification For Comparison

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Untreated Outcome Test</td>
<td>Always Take Average Treatment Effect Lower Bound</td>
<td>Local Average Treatment Effect LATE</td>
</tr>
<tr>
<td>Breast cancer incidence</td>
<td>19,505</td>
<td>-301 [0.003]</td>
<td>206 (59)</td>
<td>58 (38)</td>
</tr>
</tbody>
</table>
Breast Cancer Incidence 20 Years After Enrollment (per 10,000)

- Treatment effect homogeneity test:
  \[-301 \times (206 - 58) = -44,311 < 0\]
  \{0.023\}

- Selection homogeneity test:
  \[366 - 667 = -301 \neq 0\]
  \{0.003\}

- LATE: 58

- Always Takers: 206 (59)
- Compliers: 366
- Never Takers: 667

\[p_C = 0.19, p_I = 0.95\]
Results Are Robust Along Many Dimensions

• Alternative outcome
  – All-cause mortality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
# Robust to Alternative Outcome

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Untreated</td>
<td>Always Taker</td>
<td>Local Average</td>
<td>Test Rejects</td>
</tr>
<tr>
<td></td>
<td>Outcome Test</td>
<td>Average</td>
<td>Treatment</td>
<td>Treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treatment</td>
<td>Effect</td>
<td>Effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect</td>
<td>LATE</td>
<td>Homogeneity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td></td>
<td>(1)*((2)-(3))&lt;0</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer incidence</td>
<td>19,505</td>
<td>-301</td>
<td>206</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.003]</td>
<td>(59)</td>
<td>(38)</td>
</tr>
<tr>
<td>Alternative Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>19,505</td>
<td>-562</td>
<td>22</td>
<td>-13</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>(55)</td>
<td>(38)</td>
<td>{0.290}</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• Alternative outcome
  – All-cause mortality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
## Robust to Alternative Sample Restrictions

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Untreated Outcome Test</td>
<td>Always Taker Average Treatment Effect Lower Bound</td>
<td>Local Average Treatment Effect LATE</td>
</tr>
<tr>
<td><strong>Main Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer incidence</td>
<td>19,505</td>
<td>-301</td>
<td>206</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.003]</td>
<td>(59)</td>
<td>(38)</td>
</tr>
<tr>
<td><strong>Alternative Sample Restrictions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All excluded participants aged 40-49 at enrollment</td>
<td>30,925</td>
<td>-1,237</td>
<td>309</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>(45)</td>
<td>(44)</td>
</tr>
<tr>
<td>All participants aged 40-49 at enrollment</td>
<td>50,430</td>
<td>-826</td>
<td>298</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>(36)</td>
<td>(30)</td>
</tr>
<tr>
<td>All participants aged 50-59 at enrollment</td>
<td>39,405</td>
<td>-1,555</td>
<td>419</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>(53)</td>
<td>(34)</td>
</tr>
<tr>
<td>All participants</td>
<td>89,835</td>
<td>-1,156</td>
<td>332</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>(30)</td>
<td>(22)</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

• Alternative outcome
  – All-cause mortality

• Alternative sample restrictions
  – Excluded participants aged 40-49
  – Aged 40-49 at enrollment
  – Aged 50-59 at enrollment
  – All participants

• Alternative definitions of mammography
  – Narrower

• Alternative follow-up lengths
### Robust to Alternative Definitions of Mammography

<table>
<thead>
<tr>
<th></th>
<th>(1) Untreated Outcome Test</th>
<th>(2) Always Taker Treatment Average Effect Lower Bound</th>
<th>(3) Local Average Treatment Effect LATE</th>
<th>(4) Test Rejects Treatment Effect Homogeneity (1)*(2)-(3)&lt;0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Specification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer incidence</td>
<td>19,505</td>
<td>-301</td>
<td>206</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>[0.003]</td>
<td>(59)</td>
<td>(38)</td>
<td>{0.023}</td>
</tr>
<tr>
<td><strong>Alternative Definitions of Mammography</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least two active study period years after enrollment</td>
<td>19,505</td>
<td>-341</td>
<td>239</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>(90)</td>
<td>(35)</td>
<td>{0.019}</td>
</tr>
<tr>
<td>At least three active study period years after enrollment</td>
<td>19,505</td>
<td>-330</td>
<td>167</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>(142)</td>
<td>(36)</td>
<td>{0.206}</td>
</tr>
<tr>
<td>All active study period years after enrollment</td>
<td>19,505</td>
<td>-178</td>
<td>158</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>[0.005]</td>
<td>(181)</td>
<td>(42)</td>
<td>{0.312}</td>
</tr>
</tbody>
</table>
Results Are Robust Along Many Dimensions

- Alternative outcome
  - All-cause mortality
- Alternative sample restrictions
  - Excluded participants aged 40-49
  - Aged 40-49 at enrollment
  - Aged 50-59 at enrollment
  - All participants
- Alternative definitions of mammography
  - Narrower
- Alternative follow-up lengths
## Robust to Breast Cancer Incidence at Alternative Follow-Up Lengths: 11-20

<table>
<thead>
<tr>
<th>Years Since Enrollment</th>
<th>N</th>
<th>(1) Untreated Outcome Test Rejects Selection Homogeneity</th>
<th>Always Taker Average Treatment Effect Lower Bound</th>
<th>(3) Local Average Treatment Effect LATE</th>
<th>(4) Test Rejects Treatment Effect Homogeneity (1)*(2)-(3)&lt;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main specification: 20</td>
<td>19,505</td>
<td>-301</td>
<td>206</td>
<td>58</td>
<td>-44,311</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.003]</td>
<td>[59]</td>
<td>[38]</td>
<td>{0.023}</td>
</tr>
<tr>
<td>19</td>
<td>19,505</td>
<td>-269</td>
<td>196</td>
<td>52</td>
<td>-38,565</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.013]</td>
<td>[58]</td>
<td>[37]</td>
<td>{0.023}</td>
</tr>
<tr>
<td>18</td>
<td>19,505</td>
<td>-311</td>
<td>210</td>
<td>54</td>
<td>-48,503</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[56]</td>
<td>[35]</td>
<td>{0.010}</td>
</tr>
<tr>
<td>17</td>
<td>19,505</td>
<td>-322</td>
<td>214</td>
<td>49</td>
<td>-52,975</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[55]</td>
<td>[34]</td>
<td>{0.005}</td>
</tr>
<tr>
<td>16</td>
<td>19,505</td>
<td>-342</td>
<td>232</td>
<td>56</td>
<td>-60,245</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[54]</td>
<td>[32]</td>
<td>{0.003}</td>
</tr>
<tr>
<td>15</td>
<td>19,505</td>
<td>-381</td>
<td>211</td>
<td>84</td>
<td>-48,650</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[50]</td>
<td>[31]</td>
<td>{0.015}</td>
</tr>
<tr>
<td>14</td>
<td>19,505</td>
<td>-404</td>
<td>201</td>
<td>80</td>
<td>-49,046</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[49]</td>
<td>[29]</td>
<td>{0.020}</td>
</tr>
<tr>
<td>13</td>
<td>19,505</td>
<td>-431</td>
<td>223</td>
<td>75</td>
<td>-63,808</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[48]</td>
<td>[28]</td>
<td>{0.007}</td>
</tr>
<tr>
<td>12</td>
<td>19,505</td>
<td>-443</td>
<td>191</td>
<td>64</td>
<td>-56,156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[44]</td>
<td>[27]</td>
<td>{0.010}</td>
</tr>
<tr>
<td>11</td>
<td>19,505</td>
<td>-423</td>
<td>195</td>
<td>55</td>
<td>-59,084</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.000]</td>
<td>[43]</td>
<td>[25]</td>
<td>{0.004}</td>
</tr>
</tbody>
</table>
Robust to Breast Cancer Incidence at Alternative Follow-Up Lengths: 1-10

<table>
<thead>
<tr>
<th>Years Since Enrollment</th>
<th>N</th>
<th>(1) Untreated Outcome Test Rejects Selection Homogeneity</th>
<th>(2) Always Taker Average Treatment Effect Lower Bound</th>
<th>(3) Local Average Treatment Effect LATE</th>
<th>(4) Test Rejects Treatment Effect Homogeneity (1)*((2)-(3))&lt;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 19,505</td>
<td>-419 [0.000]</td>
<td>200 (42)</td>
<td>47 (23)</td>
<td>-64,017 {0.000}</td>
<td></td>
</tr>
<tr>
<td>9 19,505</td>
<td>-413 [0.000]</td>
<td>192 (40)</td>
<td>34 (22)</td>
<td>-64,955 {0.000}</td>
<td></td>
</tr>
<tr>
<td>8 19,505</td>
<td>-409 [0.000]</td>
<td>175 (37)</td>
<td>35 (21)</td>
<td>-57,386 {0.000}</td>
<td></td>
</tr>
<tr>
<td>7 19,505</td>
<td>-393 [0.000]</td>
<td>177 (35)</td>
<td>46 (18)</td>
<td>-51,740 {0.000}</td>
<td></td>
</tr>
<tr>
<td>6 19,505</td>
<td>-412 [0.000]</td>
<td>185 (33)</td>
<td>50 (17)</td>
<td>-55,761 {0.000}</td>
<td></td>
</tr>
<tr>
<td>5 19,505</td>
<td>-382 [0.000]</td>
<td>180 (32)</td>
<td>45 (15)</td>
<td>-51,581 {0.000}</td>
<td></td>
</tr>
<tr>
<td>4 19,505</td>
<td>-393 [0.000]</td>
<td>152 (29)</td>
<td>46 (13)</td>
<td>-41,568 {0.003}</td>
<td></td>
</tr>
<tr>
<td>3 19,505</td>
<td>-354 [0.000]</td>
<td>104 (23)</td>
<td>37 (11)</td>
<td>-23,679 {0.012}</td>
<td></td>
</tr>
<tr>
<td>2 19,505</td>
<td>-337 [0.000]</td>
<td>63 (18)</td>
<td>25 (9)</td>
<td>-12,632 {0.030}</td>
<td></td>
</tr>
<tr>
<td>1 19,505</td>
<td>-342 [0.000]</td>
<td>35 (11)</td>
<td>20 (6)</td>
<td>-5,194 {0.097}</td>
<td></td>
</tr>
</tbody>
</table>
# 2016 USPSTF Guidelines Based on RCT’s

**Annals of Internal Medicine Review**

*Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation*  
Heidi D. Nelson, MD, MPH; Rochelle Fu, PhD; Amy Cantor, MD, MPH; Miranda Pappas, MA; Monica Dauges, BA; and Linda Humphrey, MD, MPH

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women aged 39–49 y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST II</td>
<td>11.2</td>
<td>0.64 (0.39–1.06)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.73 (0.37–1.41)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Östergötland</td>
<td>12.5</td>
<td>1.02 (0.52–1.99)</td>
</tr>
<tr>
<td>Moss et al, 2015 (27)</td>
<td>Age</td>
<td>17.5</td>
<td>0.93 (0.80–1.09)</td>
</tr>
<tr>
<td>Bjurstem et al, 2003 (25)</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.69 (0.45–1.05)</td>
</tr>
<tr>
<td>Habbema et al, 1986 (29)</td>
<td>HIP</td>
<td>14.0</td>
<td>0.75 (0.53–1.05)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>Stockholm</td>
<td>14.3</td>
<td>1.52 (0.80–2.88)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST I</td>
<td>18.2</td>
<td>0.74 (0.42–1.29)</td>
</tr>
<tr>
<td>Miller et al, 2014 (15)</td>
<td>CNBSS-1</td>
<td>21.9</td>
<td>1.04 (0.87–1.24)</td>
</tr>
<tr>
<td>Overall (I² = 25%; P = 0.230)</td>
<td></td>
<td></td>
<td>0.92 (0.75–1.02)</td>
</tr>
</tbody>
</table>
**CNBSS Consistent with Meta-analysis of RCT’s**

<table>
<thead>
<tr>
<th>Author, Year (Reference)</th>
<th>Trial Name</th>
<th>Mean Follow-up, y</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women aged 39–49 y</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST II</td>
<td>11.2</td>
<td>0.64 (0.39–1.06)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Kopparberg</td>
<td>12.5</td>
<td>0.73 (0.37–1.41)</td>
</tr>
<tr>
<td>Tabár et al, 1995 (26)</td>
<td>Östergötland</td>
<td>12.5</td>
<td>1.02 (0.52–1.99)</td>
</tr>
<tr>
<td>Moss et al, 2015 (27)</td>
<td>Age</td>
<td>17.5</td>
<td>0.93 (0.80–1.09)</td>
</tr>
<tr>
<td>Bjurstam et al, 2003 (25)</td>
<td>Gothenburg</td>
<td>13.8</td>
<td>0.69 (0.45–1.05)</td>
</tr>
<tr>
<td>Habbema et al, 1986 (29)</td>
<td>HIP</td>
<td>14.0</td>
<td>0.75 (0.53–1.05)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>Stockholm</td>
<td>14.3</td>
<td>1.52 (0.80–2.88)</td>
</tr>
<tr>
<td>Nyström et al, 2002 (30)*</td>
<td>MMST I</td>
<td>18.2</td>
<td>0.74 (0.42–1.29)</td>
</tr>
<tr>
<td><strong>Miller et al, 2014 (15)</strong></td>
<td>CNBSS-1</td>
<td>21.9</td>
<td>1.04 (0.87–1.24)</td>
</tr>
</tbody>
</table>

Overall ($I^2 = 25\%; P = 0.230$)  

0.92 (0.75–1.02)
CNBSS Protocols Varied by Age

• Patients aged 40-49:
  – Intervention group: mammography + physical examination each year for 4-5 years, then return to usual care
  – Control group: usual care

• Patients aged 50-59:
  – Intervention group: mammography + physical examination each year for 4-5 years, then return to usual care
  – Control group: physical examination each year for 4-5 years, then return to usual care
USPSTF Recommendations Differ for Women in 40’s and 50’s

• The U.S. Preventive Services Task Force (USPSTF) Assigns “grades”
  – “A” and “B” grades fully-covered under ACA

• Different grades for 40’s and 50+ (Siu, 2016)
  – “The decision to start screening mammography in women prior to age 50 years should be an individual one. (Grade C recommendation)”
  – “The USPSTF recommends biennial screening mammography for women aged 50 to 74 years. (Grade B recommendation)”
Never Takers Die More Than Compliers

- Women more likely to receive mammograms are healthier
- Breast cancer mortality without screening (Kim and Lee, 2017)