STATISTICAL STRATEGIES FOR ASSESSING THE IMPACT OF NEW HEALTH TECHNOLOGIES

Sharon-Lise T. Normand[a]

Harvard Medical School and Harvard School of Public Health

OUTLINE OF TALK

- Fundamental Problem.
- Motivating Problem - why important?
- Statistical Issues.
- Methods and Illustration.
- Concluding Remarks.

---

[a] Joint work with: Laura Mauri MD; Mass-DAC(Treacy Silbaugh BSc, Ann Lovett RN, Robert Wolf MS, Matthew Cioffi, and Katya Zelevensky BA). Thanks to Paul Dreyer PhD; Massachusetts' cardiac interventionalists and data managers.

FUNDAMENTAL SCIENTIFIC PROBLEM

Cannot efficiently produce **information** under current paradigm.

1. Technology and innovation evolve **rapidly**.
   (a) Medical devices are getting smaller & smarter.
   (b) Medical devices are providing more information.
   (c) Medical devices are more convenient for the patient.
2. RCTs are **smaller** and increasingly **not generalizable**.
   (a) Representative of about 10% of potential population.
   (b) Multiple illnesses.
   (c) Multiple protocols.
   (d) In US, lack of insurance to get subjects to participating centers.

MOTIVATING PROBLEM - CORONARY STENTING

1. Drug-eluting stents (DES) reduce the need for repeat revascularization but their long term safety relative to bare metal stents (BMS) remains uncertain.
   (a) One observational study raised concerns that, beyond the first year of treatment, DES may be associated with higher rates of thrombosis and mortality.
   (b) RCTs had limited power to detect small differences in mortality or other adverse events and are limited to healthy population compareds to those who are currently treated with coronary stenting.
   (c) $>$ 6 million DES already implanted; $>$ 1 million new DES annually; and 90% US penetration with $>$ 70% off-label.

Mauri et al. 2007 AHA; Circulation (in press).

STATISTICAL ISSUES

Want to make a **causal** statement – do DES increase the risk of death compared to BMS?

1. Predictive inference focuses on comparing outcomes between **groups of individuals** having received different treatments.
2. Causal inference is a special case in which subjects **who could have** received either treatments are identified and used to infer treatment effects.
   (a) Want to know what **would have** happened to a subject's survival had the patient received a **different treatment** than the one observed.
3. Causal inference involves a **counterfactual** outcome (not observed).

NOTATION

$$
\begin{aligned}
Y &= \text{observed response (e.g., death)} \\
Y(1) &= \text{experimental outcome (implanted with DES)} \\
Y(0) &= \text{non-experimental outcome (implanted with BMS)} \\
Z &= \text{binary treatment indicator (0=BMS; 1=DES)} \\
X &= \text{covariates} \\
N &= \text{total number of subjects} \\
\theta &= Y(1) - Y(0) \ \ \text{causal effect}
\end{aligned}
$$

The **observed** response is

$$Y = Y(1)Z + Y(0)(1 - Z) \tag{1}$$

NOTATION

Causal Effect Desired:

$$\theta = E(Y(1)) - E(Y(0)) = \theta_1 - \theta_0 \tag{2}$$

causal inference

Effect Most Commonly Estimated:

$$\alpha(X) = E(Y \mid X, Z = 1) - E(Y \mid X, Z = 0) \tag{3}$$

predictive inference

ASSUMPTIONS UNDER RANDOMIZATION

1. Treatment assignment is independent of outcomes,

$$(Y(1), Y(0)) \perp Z$$

2. Subjects have a positive probability of receiving the treatment, $1 > P(Z = 1 \mid X) > 0$.
3. The treatment effect for a subject does not depend on the treatment assignments of other subjects.
   (a) Cox (1992).
   (b) Rubin (1978) - Stable Unit Treatment Value Assignment [SUTVA].

IMPLICATIONS OF RANDOMIZATION

$$\begin{aligned}
\theta(X) &= E(Y(1) - Y(0) \mid X, Z = 1) \\
&= E(Y \mid X, Z = 1) - E(Y(0) \mid X, Z = 1) \\
\text{because} &\quad (Y(1), Y(0)) \perp Z \ \ \forall X \\
&= E(Y \mid X, Z = 1) - E(Y \mid X, Z = 0) \qquad (4)
\end{aligned}$$

That is, $E(Y(0) \mid X, Z = 1) = E(Y \mid X, Z = 0)$

Selection Bias: $E(Y(0) \mid X, Z = 1) \neq E(Y \mid X, Z = 0)$

STATISTICAL ISSUES IN OBSERVATIONAL SETTING

1. Lack of **randomization**:
   (a) Experimenter does not determine the assignment of treatments to subjects.
   (b) Treatment assignment mechanism is unknown and must be estimated.
   (c) Not every subject has a **chance** of receiving the treatment.
   (d) Unmeasured confounders may be present.
   - Variables that are related to both the probability of treatment and to the outcome confound the relationship between treatment assignment and outcome.

ASSUMPTIONS TYPICALLY MADE IN AN OBSERVATIONAL SETTING

1. No unmeasured confounders conditional on X, that is,

$$(Y(1), Y(0)) \perp Z \mid X.$$

   Within subpopulations defined by X, we have random treatment assignment.

2. Positivity Condition: $1 > P(Z = 1 \mid X) > 0$. The probability needs to be bounded away from zero.

3. SUTVA

(1) & (2) imply **strong** ignorability.

REGRESSION ESTIMATOR[a].

$$\hat{\theta}^R = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{E}(Y \mid Z=1, X_i) - \hat{E}(Y \mid Z=0, X_i)\right).$$

1. Advantages
   - Can use standard software to fit.
   - Simple to interpret and familiar to multiple audiences.
2. Disadvantages:
   - Requires a parametric model; if non-parametric, then X has to be low-dimensional.
   - Comparability of treatment groups difficult.
   - Impose linearity & extrapolate over different regions of X.
   - Increase/overcorrect for bias if V(X) differs between treatment groups.

---

[a] $\hat{E}$ =parametric regression

PROPENSITY SCORE ESTIMATOR

Let $S = P(Z = 1 \mid X)$ Propensity Score

(Rosenbaum and Rubin, 1983)

$$\begin{aligned}
\theta_1 - \theta_0 &= E(Y(1) - Y(0)) \\
&= E(Y \mid Z = 1, S = s) - E\left(\frac{(Y)(1-Z)}{P(Z = 0 \mid S = s)}\right) \\
&= \cdots \\
&= E(Y \mid Z = 1, S = s) - \sum W(i,j)(Y \mid Z = 0)
\end{aligned}$$

where $W(i,j)$ are weights depending on distances between i and j.

$$\hat{\theta}^M = \frac{1}{N} \sum_{i \in S*} \left\{ Y_i Z_i - \sum_{j \in S*} W(i,j) Y_i \right\}$$

and $S*$ is a region of common support that is measured on the basis of S_i. (D'Agostino Jr.,1998).

MATCHING ESTIMATORS

Advantages

- Comparability of treatment groups apparent.
- Analyses are simple (paired-t-tests; McNemar's test).
- Variance in average treatment effect is smaller in matched samples.

Disadvantages

- Need a large control group (inexact matches).
- Could be a loss in efficiency.
- Need to get propensity score model right.

MATCHING ESTIMATOR

Gu and Rosenbaum (1993)

1. Structure of Matches:
   (a) k-k matches.
   (b) exact matching.
2. Assignment of Matches:
   (a) nearest available match, full matching, optimal matching.
3. Closeness of Matches: $||X_i - X_j|| \;\; j \in \{Z = 0\}; \;\; i \in \{Z = 1\}$:
   (a) any match: find a X_j such that $\min ||X_i - X_j||$.
   (b) restricted: find a X_j such that $||X_i - X_j|| < \epsilon$ for small ϵ.

ASESSING POSITIVITY CONDITION

How to assess extent of overlap in observed confounders between the two treatment groups?

- Standardized Differences: a measure of non-overlap (sociology).

$$s_d = \left(\frac{\bar{x}_1 - \bar{x}_0}{s_p} \right) \times 100. \tag{5}$$

- s_p is the pooled standard deviation.
- If $\mid s_d \mid < 10\%$, OK (Cohen's Effect Size).
  - Percent non-overlap between the two treatment groups;
  - Indicates $\approx 7.7\%$ of the area covered by both groups combined is non-overlapped.
- Iterative process.
- Typically better than examining test statistics.

UNMEASURED CONFOUNDERS

$$\log\left(\frac{1}{\Gamma}\right) \leq \log\frac{P(Z=1 \mid X=x, U=u)}{P(Z=0 \mid X=x, U=u)} = \beta(X) + \beta_1 u \leq \log(\Gamma)$$

- $\beta_1 = 0$ then treatment assignment is strongly ignorable
- $\beta_1 \neq 0$ then odds differ by at most $\exp(\beta_1(u - u*))$
- If $\Gamma = 2$ then two subjects who appear similar on the same covariates, could differ in their odds of receiving a DES by as much as a factor of 2.
  - One patient could be twice as likely as the other to get the DES
- Can calculate bounds on p-values for statistical test (Rosenbaum, 1987).
- A study is **insensitive** if extreme values of Γ are required to change conclusions.

UNMEASURED CONFOUNDERS

Bounds on p-values for McNemar's test:

$$\sum_{a=y_d}^{D} \binom{D}{a} (p^-)^a (1-p^-)^{D-a} \le \text{prob}\{T \ge y_d\} \le \sum_{a=y_d}^{D} \binom{D}{a} (p^-)^a (1-p^-)^{D-a}$$

$$p^- = \frac{\Gamma}{1+\Gamma} \text{ and } p^+ = \frac{1}{1+\Gamma}$$

- D is number of discordant pairs; y_d is number BMS dying.
- T is the sign score statistic.
- When $\beta_1 = 0$, $p^- = p^+ = \frac{1}{2}$, get usual significance level.

DRUG-ELUTING AND BARE METAL STENTING

- All adults undergoing PCI with stenting between April 1, 2003 and September 30, 2004 at all non-federal acute care hospitals in Massachusetts, USA.
- State-mandated database:
  - Clinical and procedural factors collected prospectively.
  - Selected data elements adjudicated by interventionalists.
  - Linked to hospital billing data; to vital statistics; and to clinical and procedure factors in a CABG database also mandated by the state.
- Study procedures, investigators, and data security procedures reviewed annually by the Harvard Medical School Internal Review Board.

DRUG-ELUTING AND BARE METAL STENTING

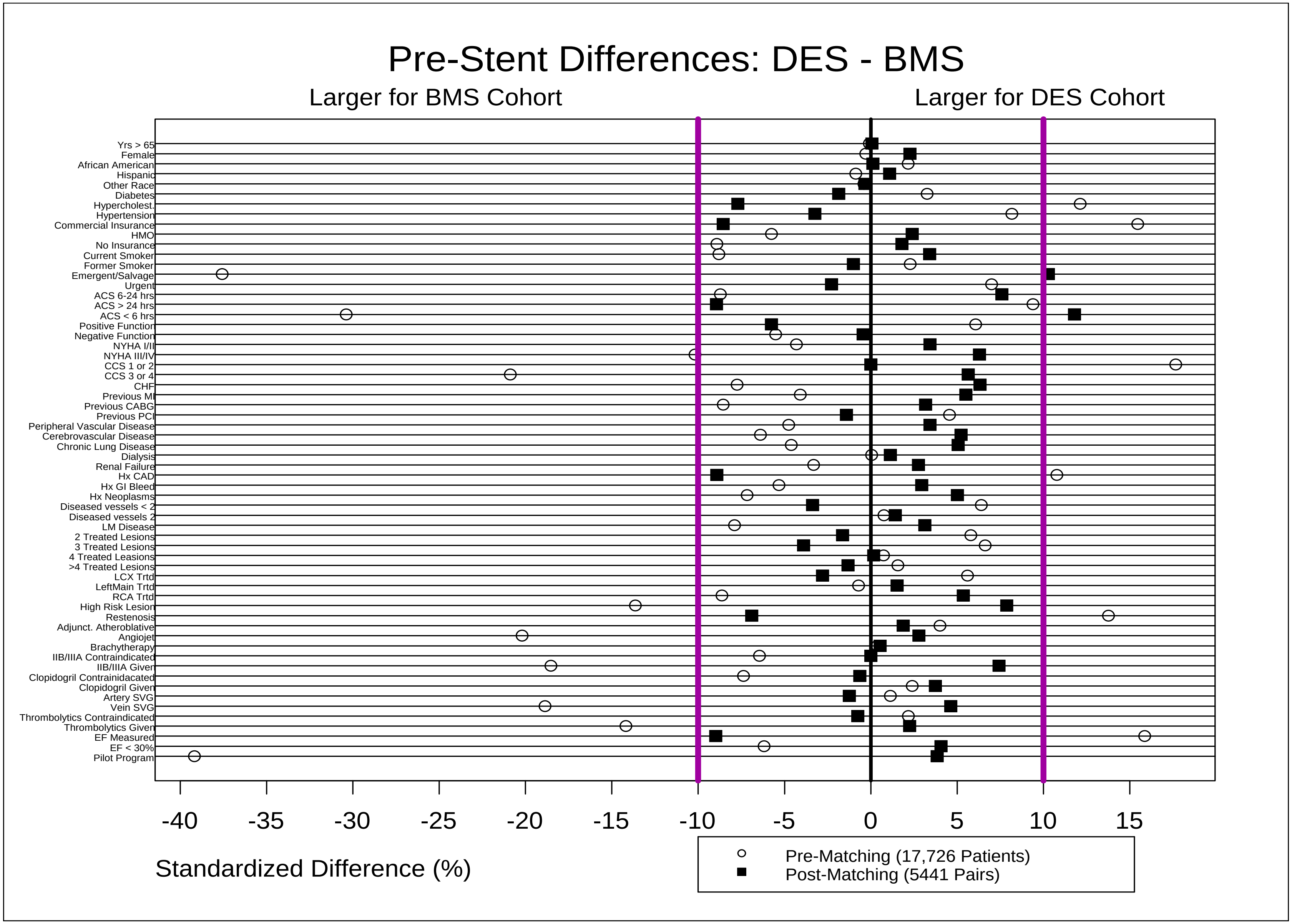
- Treatment Groups: **DES** if all stents used during the index admission were DES versus **BMS** if all stents were BMS.
  - Subjects implanted with both types of stents during the index admission were eliminated.
- Primary outcome: **death** from any cause within 2 years of the index procedure.
  - 2-year complete follow-up on all subjects.
- Secondary outcomes:
  - Myocardial infarction within 2-years.
  - Repeat revascularization within 2-years defined as any CABG or PCI following index procedure.

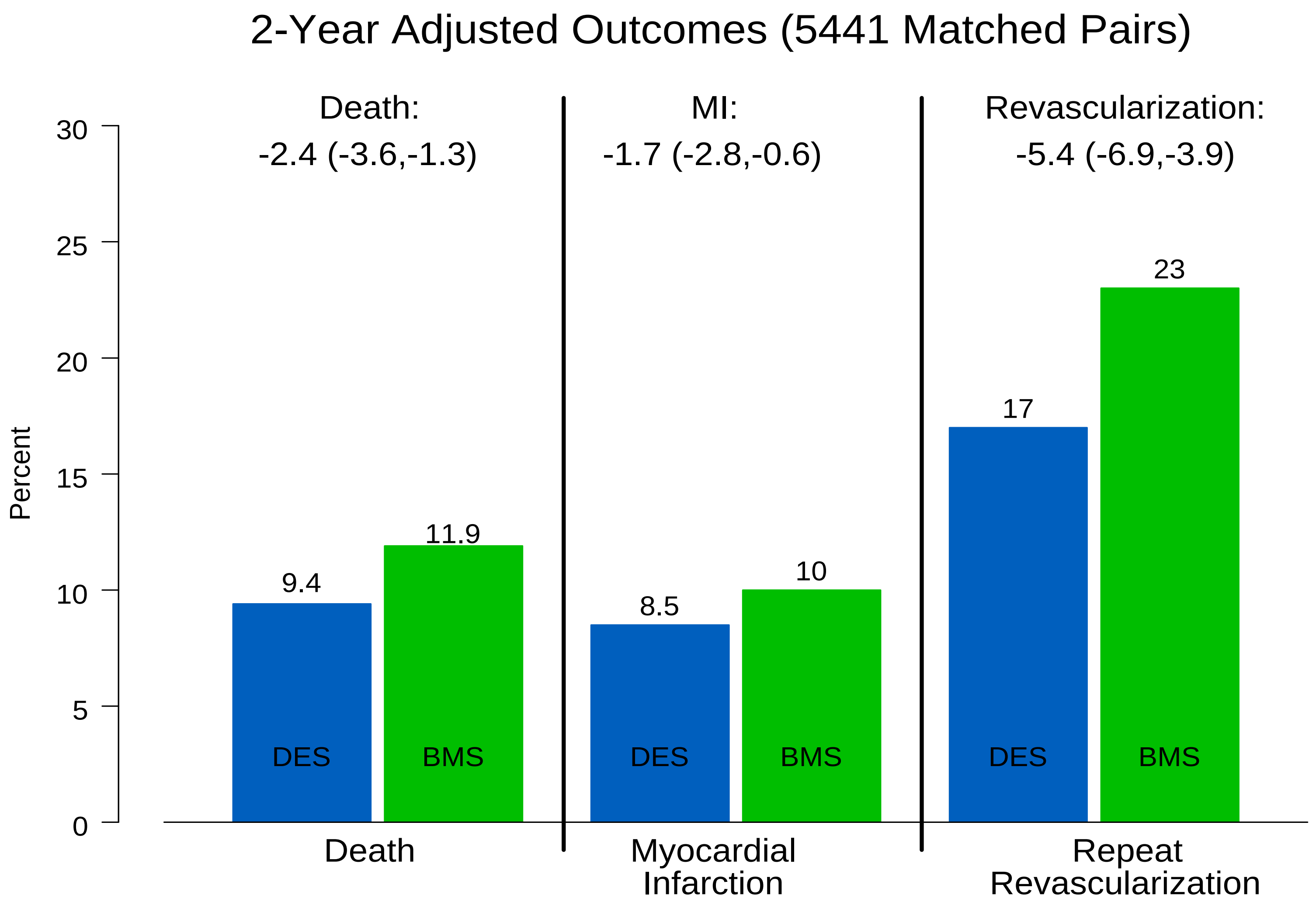
DRUG-ELUTING AND BARE METAL STENTING

- N = 21,019 PCI patients in Massachusetts undergoing PCI with stenting between April 1, 2003 and September 30, 2004.
  - Excluded 1538 (7.3%) non-MA residents.
  - Excluded 564 (2.7%) PCI patients unlinkable to billing data.
  - Excluded 1191 (5.7%)patients implanted with both DES and BMS during index admission.
- Final Cohort = **17,726** (84.3%) patients:
  - 11,516 patients with DES only (65%)
  - 6,210 with BMS only (35%)

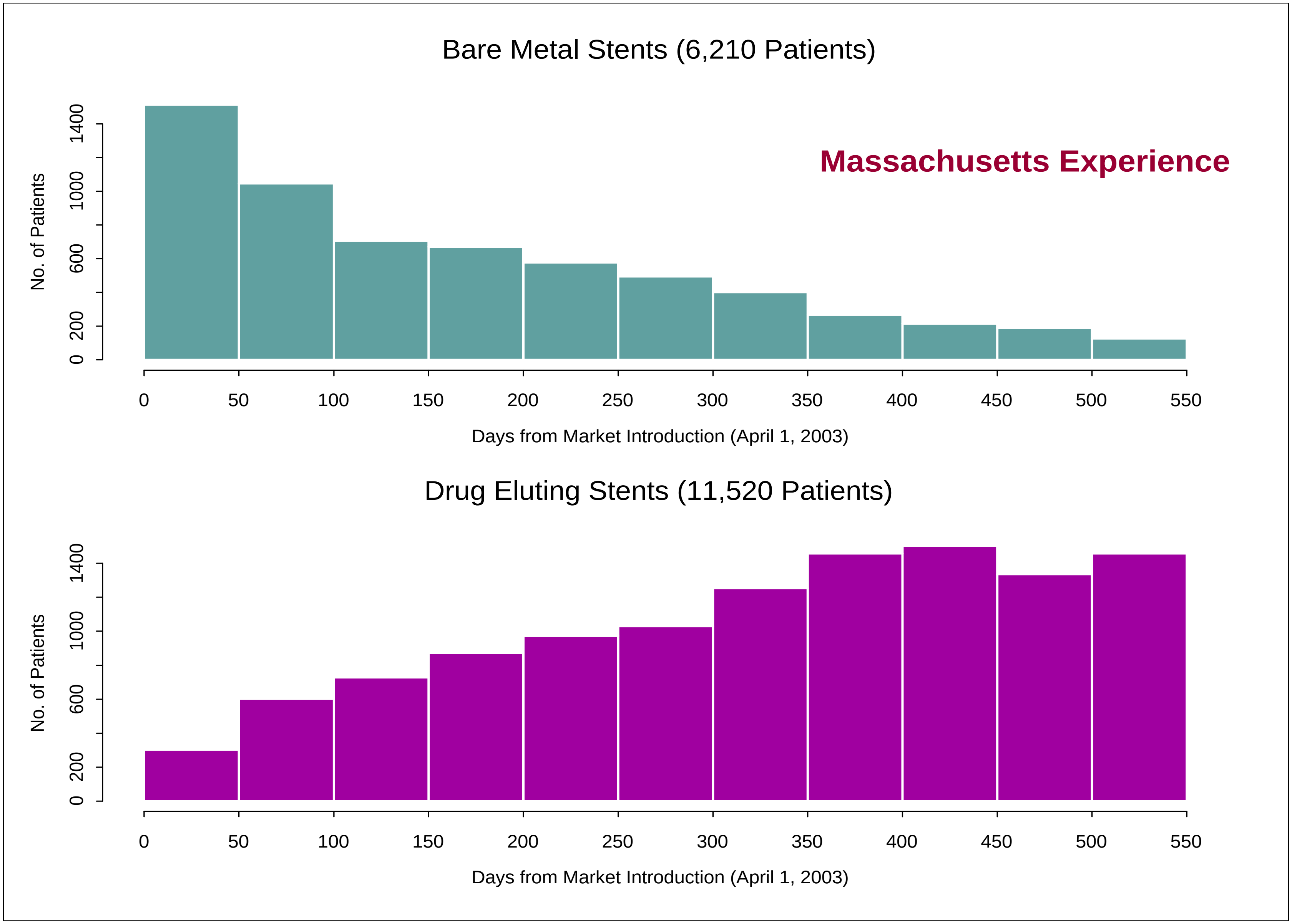
DES EXAMPLE

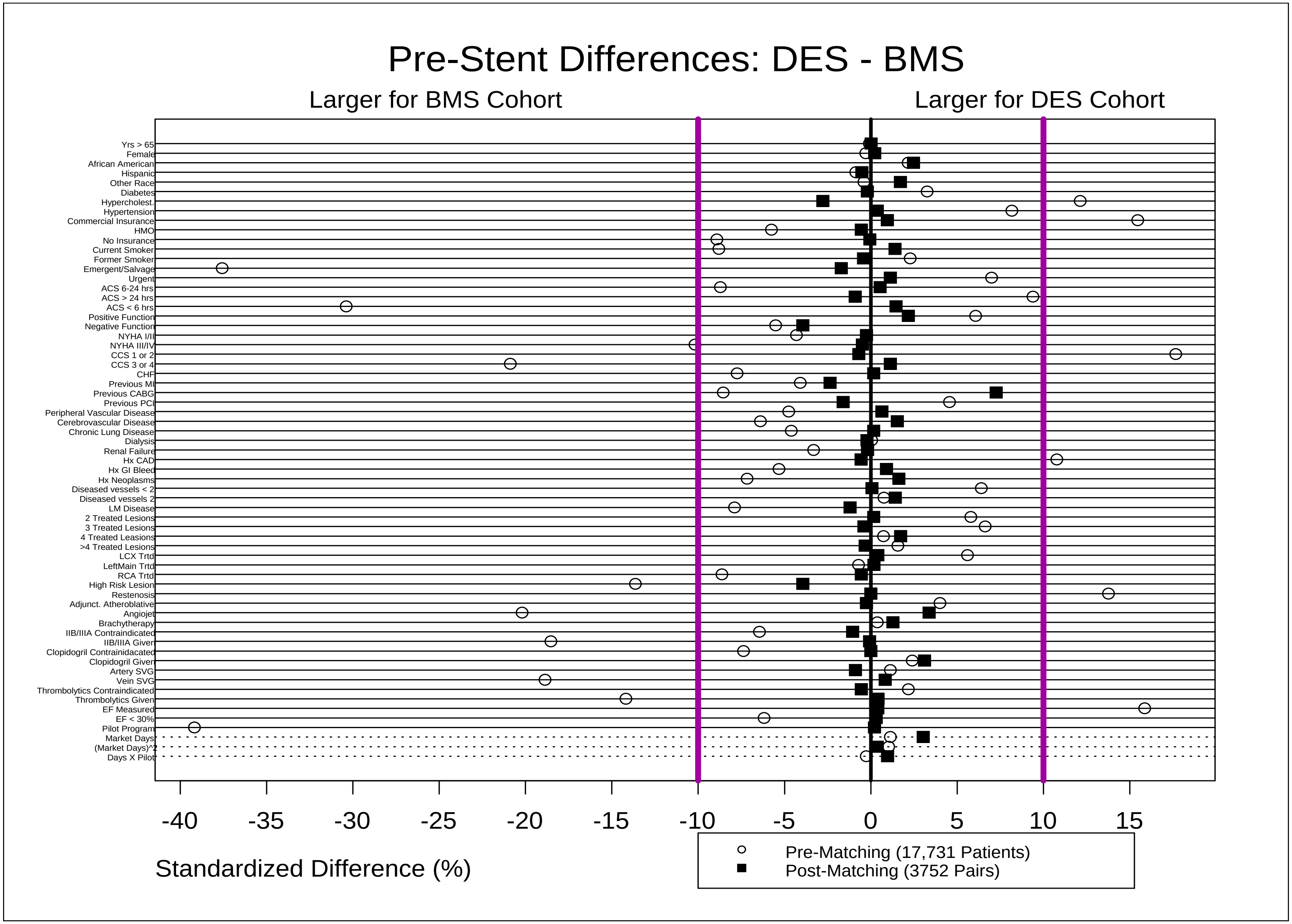
1. Propensity score model included **63** variables (demographic, comorbidity, presenting risk, procedural, and lesion characteristics).
2. Matching: 1-1 matches using nearest available match.
3. Closeness of Matches: $||X_i - X_j|| \;\; j \in \{Z = 0\}; \;\; i \in \{Z = 1\}$ restricted using $\epsilon =$ **0.60** $\times \sqrt{\text{var}(\log(S_i))}$
4. Paired t-tests for outcome differences.
5. Unmeasured confounding:
   (a) Examined **2-day** differences in outcomes.
   (b) **Bounds** in McNemar's test (paired-t-test).


Pre-Stent Differences: DES - BMS
Larger for BMS Cohort
Larger for DES Cohort
Yrs > 65
Female
African American
Hispanic
Other Race
Diabetes
Hypercholest.
Hypertension
Commercial Insurance
HMO
No Insurance
Current Smoker
Former Smoker
Emergent/Salvage
Urgent
ACS 6-24 hrs
ACS > 24 hrs
ACS < 6 hrs
Positive Function
Negative Function
NYHA I/II
NYHA III/IV
CCS 1 or 2
CCS 3 or 4
CHF
Previous MI
Previous CABG
Previous PCI
Peripheral Vascular Disease
Cerebrovascular Disease
Chronic Lung Disease
Dialysis
Renal Failure
Hx CAD
Hx GI Bleed
Hx Neoplasms
Diseased vessels < 2
Diseased vessels 2
LM Disease
2 Treated Lesions
3 Treated Lesions
4 Treated Leasions
>4 Treated Lesions
LCX Trtd
LeftMain Trtd
RCA Trtd
High Risk Lesion
Restenosis
Adjunct. Atheroblative
Angiojet
Brachytherapy
IIB/IIIA Contraindicated
IIB/IIIA Given
Clopidogril Contrainidacated
Clopidogril Given
Artery SVG
Vein SVG
Thrombolytics Contraindicated
Thrombolytics Given
EF Measured
EF < 30%
Pilot Program
-40
-35
-30
-25
-20
-15
-10
-5
0
5
10
15
Standardized Difference (%)
Pre-Matching (17,726 Patients)
Post-Matching (5441 Pairs)



2-Year Adjusted Outcomes (5441 Matched Pairs)
Death:
-2.4 (-3.6,-1.3)
MI:
-1.7 (-2.8,-0.6)
Revascularization:
-5.4 (-6.9,-3.9)
Percent
30
25
20
15
10
5
0
9.4
11.9
8.5
10
17
23
DES
BMS
DES
BMS
DES
BMS
Death
Myocardial Infarction
Repeat Revascularization


Bare Metal Stents (6,210 Patients)
Massachusetts Experience
No. of Patients
0
200
600
1000
1400
0
50
100
150
200
250
300
350
400
450
500
550
Days from Market Introduction (April 1, 2003)
Drug Eluting Stents (11,520 Patients)
No. of Patients
0
200
600
1000
1400
0
50
100
150
200
250
300
350
400
450
500
550
Days from Market Introduction (April 1, 2003)


Pre-Stent Differences: DES - BMS
Larger for BMS Cohort
Larger for DES Cohort
Yrs > 65
Female
African American
Hispanic
Other Race
Diabetes
Hypercholest.
Hypertension
Commercial Insurance
HMO
No Insurance
Current Smoker
Former Smoker
Emergent/Salvage
Urgent
ACS 6-24 hrs
ACS > 24 hrs
ACS < 6 hrs
Positive Function
Negative Function
NYHA I/II
NYHA III/IV
CCS 1 or 2
CCS 3 or 4
CHF
Previous MI
Previous CABG
Previous PCI
Peripheral Vascular Disease
Cerebrovascular Disease
Chronic Lung Disease
Dialysis
Renal Failure
Hx CAD
Hx GI Bleed
Hx Neoplasms
Diseased vessels < 2
Diseased vessels 2
LM Disease
2 Treated Lesions
3 Treated Lesions
4 Treated Leasions
>4 Treated Lesions
LCX Trtd
LeftMain Trtd
RCA Trtd
High Risk Lesion
Restenosis
Adjunct. Atheroblative
Angiojet
Brachytherapy
IIB/IIIA Contraindicated
IIB/IIIA Given
Clopidogril Contrainidacated
Clopidogril Given
Artery SVG
Vein SVG
Thrombolytics Contraindicated
Thrombolytics Given
EF Measured
EF < 30%
Pilot Program
Market Days
(Market Days)^2
Days X Pilot
-40 -35 -30 -25 -20 -15 -10 -5 0 5 10 15
Standardized Difference (%)
Pre-Matching (17,731 Patients)
Post-Matching (3752 Pairs)

2-YEAR MORTALITY: MCNEMAR'S TEST

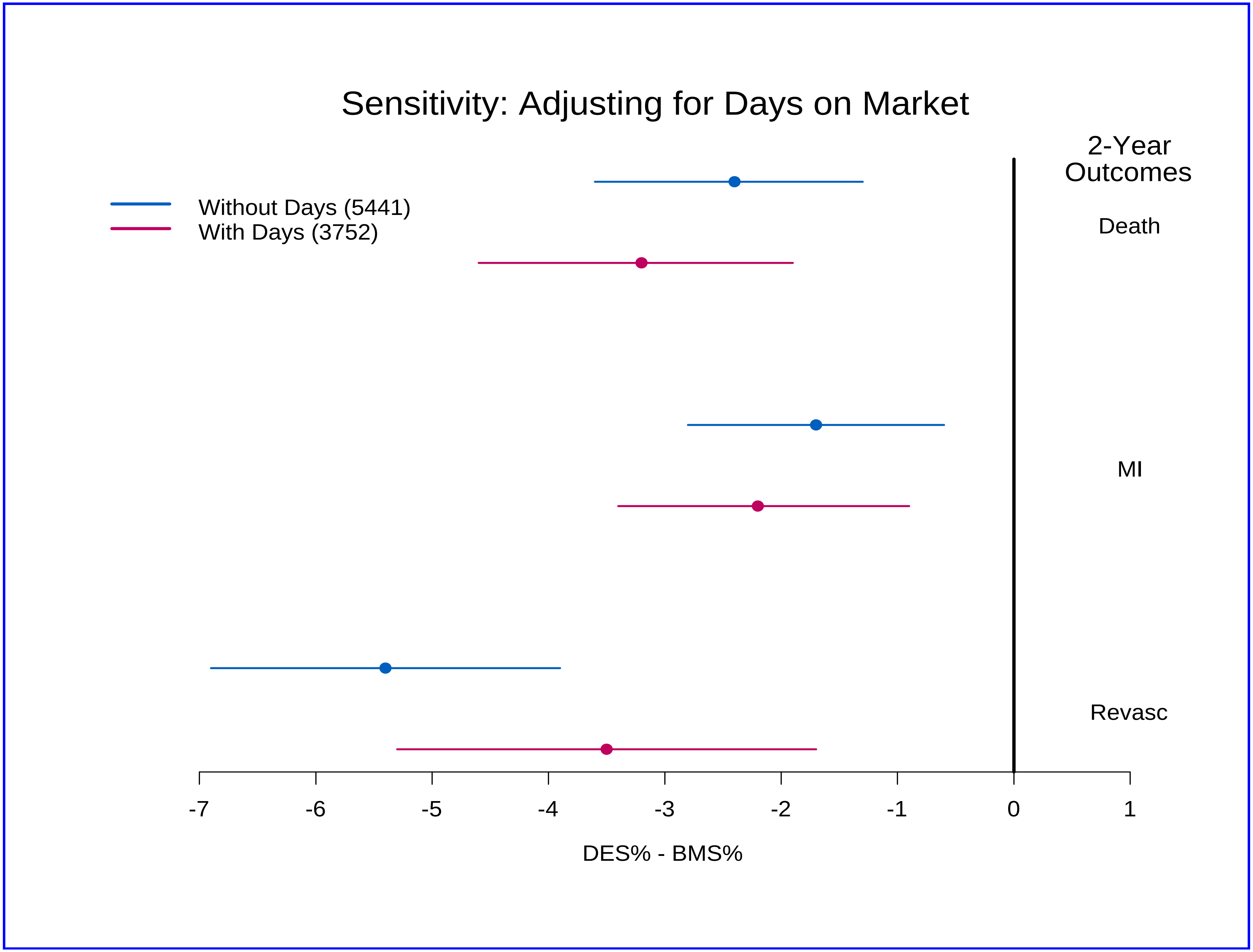
Without Days

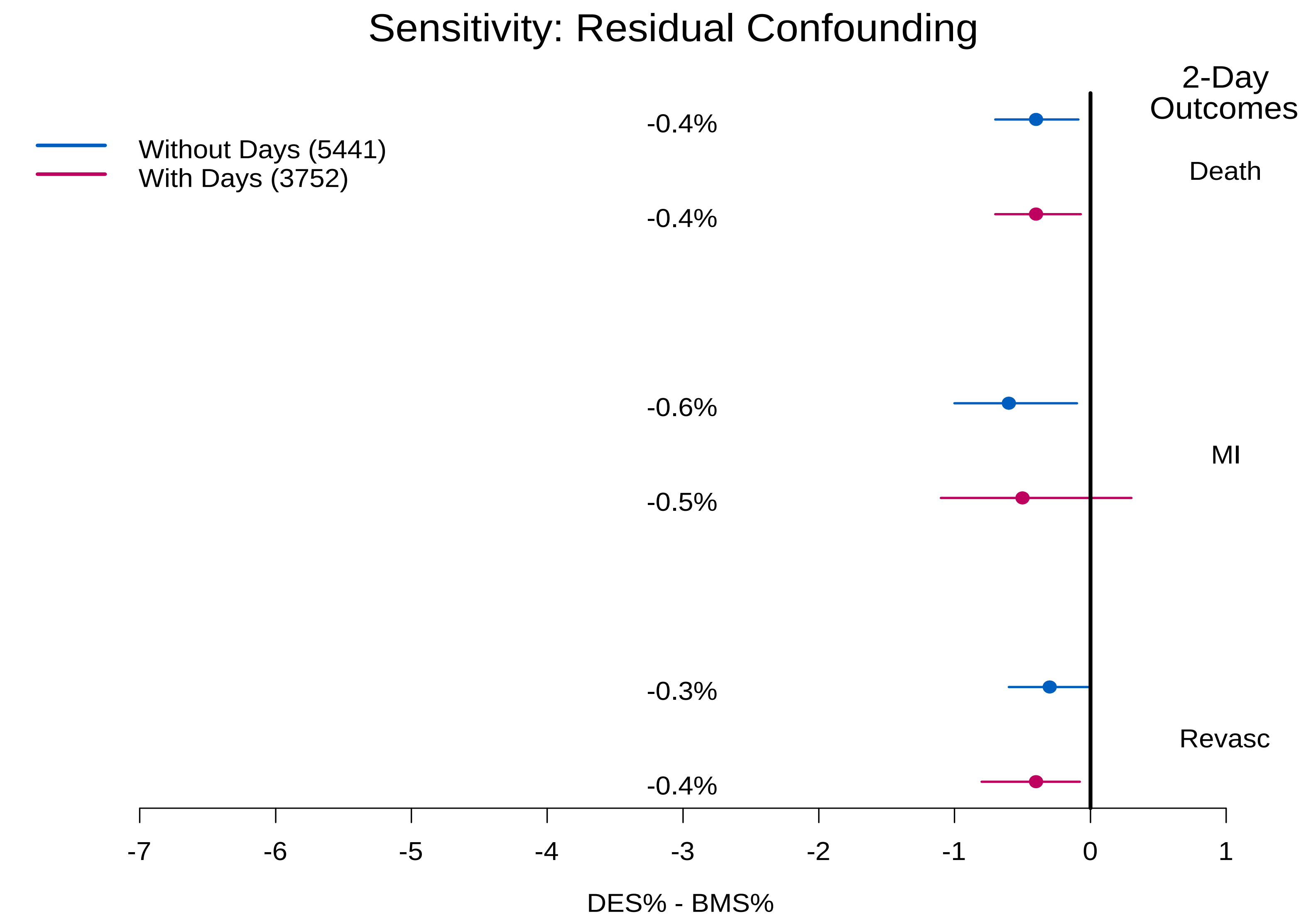
	BMS		
	Alive	Dead	Total
Alive	4349	**578**	4922
Dead	**445**	69	514
Total	4794	647	**5441**

With Days

	BMS		
	Alive	Dead	Total
Alive	3032	**402**	3434
Dead	**281**	37	318
Total	3313	439	**3752**

1. Without Days: 1023/5441 = 18.8% discordant pairs with 57% of BMS patients dying within 2 years of stent placement.
   (a) $\chi^2 = 17.03$ on 1 dof yields $p < 0.0001$.
2. With Days: 683/3752 = 18.2% discordant pairs with 58.9% of BMS patients dying within 2 years of stent placement.
   (a) $\chi^2 = 21.08$ on 1 dof yields $p < 0.0001$.


Sensitivity: Adjusting for Days on Market
2-Year Outcomes
Without Days (5441)
With Days (3752)
Death
MI
Revasc
-7
-6
-5
-4
-3
-2
-1
0
1
DES% - BMS%


Sensitivity: Residual Confounding
2-Day
Outcomes
Without Days (5441)
With Days (3752)
-0.4%
-0.4%
Death
-0.6%
-0.5%
MI
-0.3%
-0.4%
Revasc
-7
-6
-5
-4
-3
-2
-1
0
1
DES% - BMS%

SENSITIVITY ANALYSIS FOR MCNEMAR'S TEST

$$\frac{1}{\Gamma} \leq \left(\frac{P(\text{DES} \mid X, u)}{P(\text{BMS} \mid X, u^*)} \right) \leq \Gamma$$

Without Days			**With Days**		
Γ	Min P	Max P	Γ	Min P	Max P
1	2.0×10^{-5}	2.0×10^{-5}	1	2.1×10^{-6}	2.1×10^{-6}
1.1	7.0×10^{-7}	4.5×10^{-3}	1.1	2.5×10^{-9}	3.9×10^{-4}
1.2	9.2×10^{-13}	0.11	1.2	1.5×10^{-12}	1.3×10^{-2}

1. Data present range of significance levels for unmeasured confounders of various levels.
2. Model that includes days on market is less sensitive than model without days but H_0 could become plausible for some missing confounders with $\Gamma = 1.2$.

Practice Patterns in Other Studies

Geographic Region	Adult Population (millions)	Total Patients	% DES	Dates of Stenting
Sweden[a]	9	19,771	30	1/03 - 12/04
Western Denmark[b]	3	12,395	29	1/02 - 6/05
Ontario[c]	9	13,353	38	12/03 - 3/05
Massachusetts[d]	6	17,726	**65**	4/03 - 9/04

[a] New England Journal of Medicine, 2007;356:1009-1019.

[b] Journal of the American College of Cardiology, 2007;50:463.

[c] New England Journal of Medicine, 2007;357:1393-1342.

[d] American Heart Assocation Scientific Meetings, 2007; Circulation (in press).

SCAAR (NEJM 2007)

19771 Swedish patients implanted 2003 – 2004

DES penetration = 31%

RR Death: **1.18 (1.03,1.35)**

Tu (NEJM 2007)

13353 Ontario patients implanted 2003 – 2004

DES penetration = 38%

RR Death = **0.70 (p-value < 0.0001).**

Mauri (Circ 2007)

17721 MA patients implanted 2003 – 2004.

DES penetration = 65%

RR Death = **0.79 (p-value < 0.001)**

SCAAR (NEJM 2007)

19771 Swedish patients implanted 2003 – 2004

DES penetration = 31%

RR Death: **1.18 (1.03,1.35)**

Tu (NEJM 2007)

13353 Ontario patients implanted 2003 – 2004

DES penetration = 38%

RR Death = **0.70 (p-value < 0.0001).**

Mauri (Circ 2007)

17721 MA patients implanted 2003 – 2004.

DES penetration = 65%

RR Death = **0.79 (p-value < 0.001)**

SCAAR (2007)[§]

35262 Swedish patients implanted 2003 – 2005

DES penetration = 39%

RR Death: **1.03 (0.94, 1.14).**

CONCLUDING REMARKS

1. Adopted a semi-parametric approach to inference.
2. Sensitivity analyses permit formal quantification of the size of biases that could explain the association between outcome and treatment.
3. Propensity score provides a metric to perform formal assessments of population comparability:
   (a) Positivity condition.
   (b) Facilitates sensitivity analyses.
4. Extensions to longitudinal designs: more work on sensitivity analyses.
5. Extensions to more treatment groups: more work needed.

ATYPICAL ANTIPSYCHOTICS[a]

Antipsychotic therapy is widely used to treat behavioral problems in older adults with dementia but cohort studies have not fully characterized the short-term harm associated with these agents.

1. Used for short periods to treat agitation in clinical practice.
2. Of residents newly admitted to nursing homes, 17% are started on an atypical within 100 days of their admission, and 10% receive only a single prescription.
3. No RCT evidence to support short-term use.
4. Safety concerns involve extrapyramidal symptoms, falls, hip fractures, cerebrovascular events, or death.

---

[a] Joint work with: Paula Rochon MD, Sudeep Gill MD, Geoffrey Anderson MD, Kathy Sykora MSc, Loarraine Liscombe MD, Chaim Bell MD, Jerry Gurwitz MD. *Archives of Internal Medicine*, 2008;168(10).

SHORT TERM ANTIPSYCHOTIC USE

Characteristic	No Antipsychotic n = 6894	Atypical n = 6894	Conventional n = 6894
Mean Age, (SD)	81.6 (7.0)	81.9 (7.0)	81.5 (7.0)
Women	4400 (63.8)	4208 (61.0)	4187 (60.7)
Low Income	2307 (33.5)	2510 (36.4)	2507 (36.5)
Psychotropic Meds in past 120 days			
Antidepressants	1611 (23.4)	1623 (23.5)	1633 (23.7)
Anxiolytics	1703 (24.7)	2227 (32.3)	2220 (32.2)
Anticonvulsants	174 (2.5)	222 (3.2)	240 (3.5)
Antimanic agents	12 (0.2)	17 (0.2)	16 (0.2)
Hosp. Admitted for delerium	134 (1.9)	257 (3.7)	257 (3.7)