Extension of Disease Risk Score-Based Confounding Adjustments for Multiple Outcomes Of Interest- An Empirical Evaluation

Rishi J. Desai, Richard Wyss, Yinzhu Jin, Justin Bohn, Sengwee Toh, Austin Cosgrove, Adee Kennedy, Jessica Kim, Clara Kim, Rita Ouellet-Hellstrom, Sara Karami, Jacqueline M. Major, Aaron Niman, Shirley V. Wang, and Joshua J. Gagne

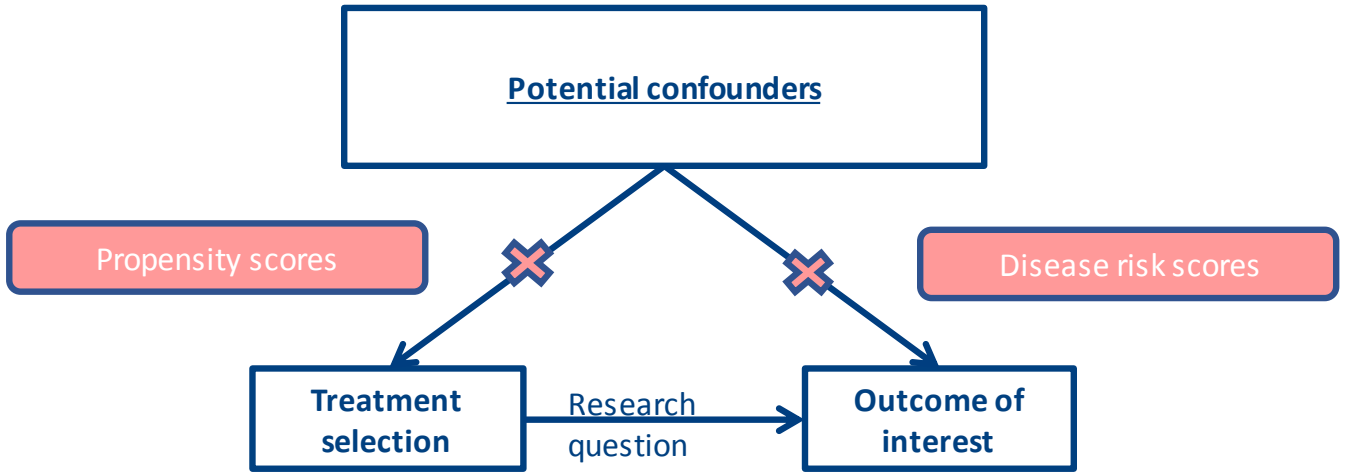
rdesai@bwh.harvard.edu @Rishidesai11

Disclosures

- The authors have no conflicts of interest to disclose
- This work was supported by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I
- This presentation reflects the views of the authors and not necessarily those of the U.S. FDA

Disease risk scores

- Disease risk scores (DRSs) are confounder summary scores used as tools for confounding control in observational studies
- Summarize multiple confounding variables into a single scalar score *based on their associations with the disease* (i.e., outcome) of interest



Study design features that influence the value or feasibility of disease risk scores (DRS) or propensity scores (PS)

Study feature or analytic goal	Impact on DRS	Impact on PS
Ample historical data (before new treatment)	Very useful for DRS development	Informs variable selection, but not generally used in estimation
Rare outcome	Greatly limits DRS development and usefulness	PS particularly valuable, but limited ability to exclude possible instrumental variables
Rare exposure	Little impact on DRS	Limits estimation of PS
Rapidly evolving treatment indications	Little impact on DRS	Challenges ability to fit PS and suggests time interactions or time- specific PS
Interest in >1 outcome/>2 exposures	DRS may be particularly useful with >2 exposures/require multiple DRS for multiple outcomes	Single PS useful for multiple outcomes with attention to risk factors for all outcomes in PS development
Interest in effect measure modification	Disease risk a natural scale for evaluation	Although less natural than the risk scale, a potentially principled summary scale
Balance disease risk across covariates	DRS a natural scale for stratification/matching	Stratification/matching on PS may provide secondary balance
Balance treatment preference across covariates	Stratification/matching on DRS may provide secondary balance	PS a natural scale for stratification/matching
Exclude (trim) subjects in one treatment group without comparable, alternatively treated comparators	Potentially valuable to exclude high or low risk subjects without comparators	Potentially valuable to exclude subjects in PS tails without comparators
Relevance of the C-statistic	A high C-statistic provides some evidence of good performance in discriminating subjects who will vs those who will not develop the outcome	A high C-statistic can indicate clearly different indications for use of the compared treatments with possibly substantial areas of non-overlap in PS distributions

Glynn et al. *PDS* 2012

Objective

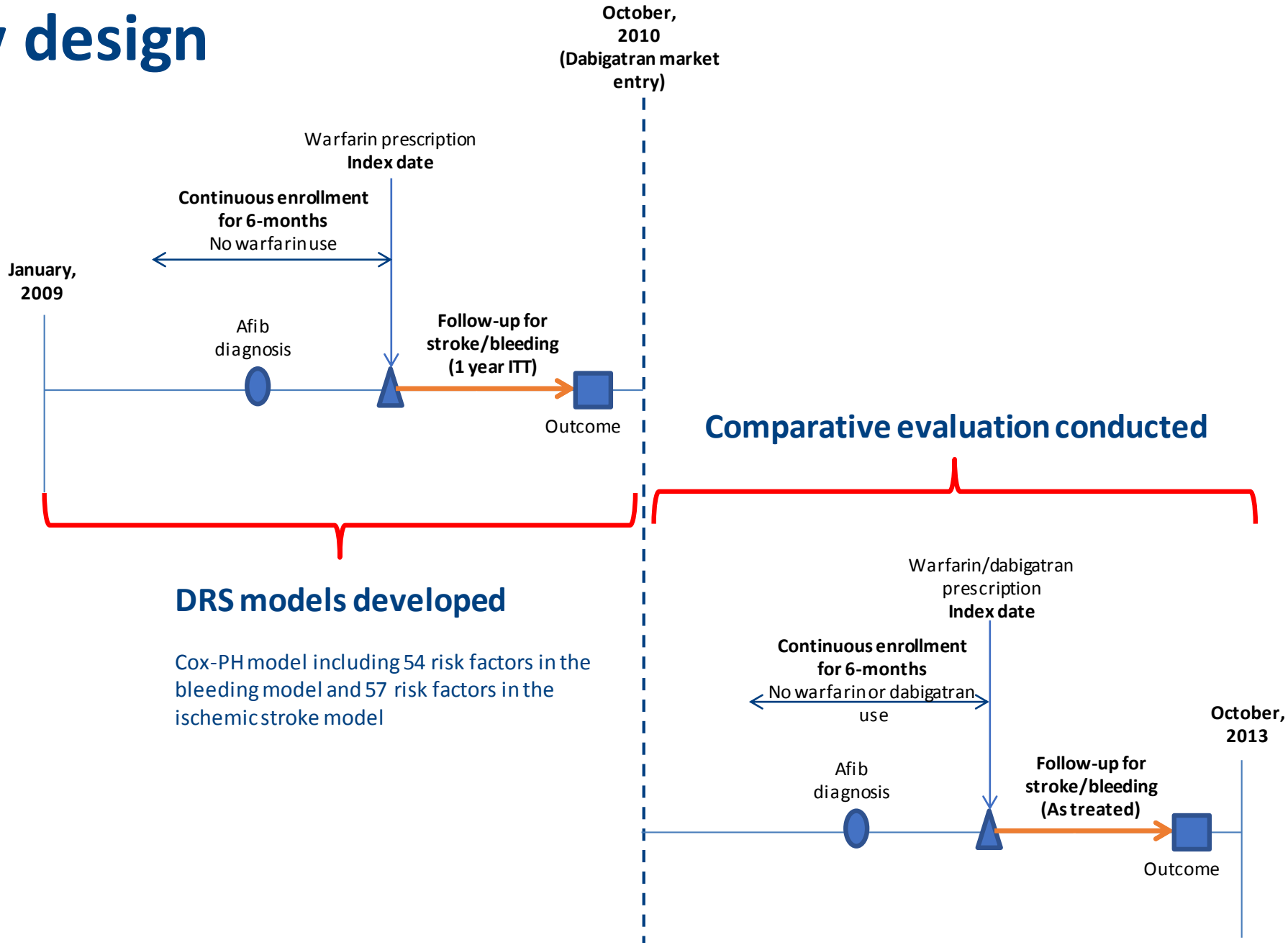
- To evaluate approaches for use of multiple disease risks scores to facilitate analyses when the interest is in comparing multiple outcomes between two treatment groups
  - Empirical example of dabigatran vs warfarin in atrial fibrillation patients on two outcomes: ischemic stroke and major bleeding

Methods

Empirical Example

- Dabigatran vs warfarin in atrial fibrillation patients on two outcomes: ischemic stroke and major bleeding
- Truven Marketscan data, 2010-2013
  - An employer-based health insurance claims database from the US
  - Longitudinal information on pharmacy claims, inpatient and outpatient visits is available

Study design



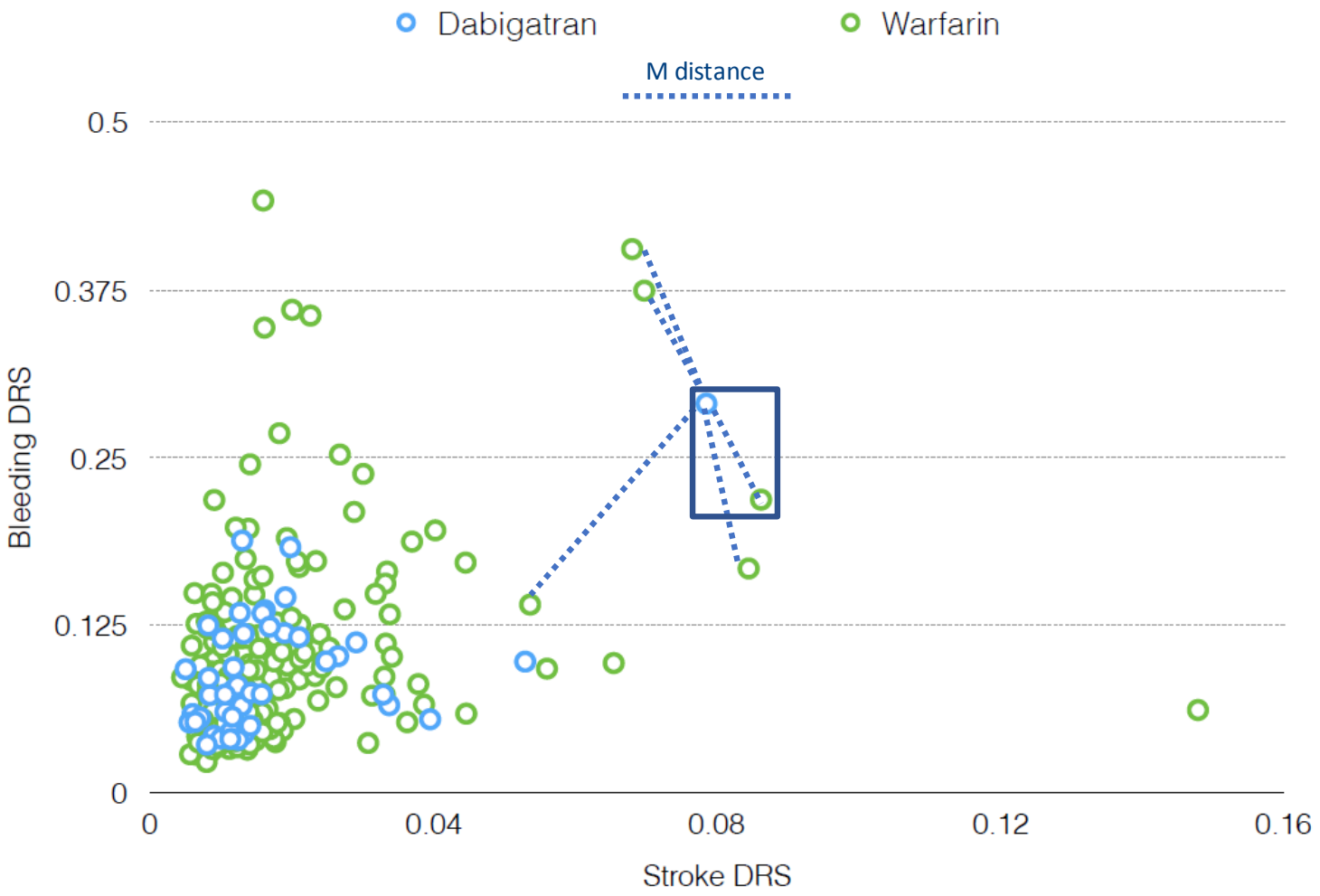
Approach 1: Prognostic propensity scores (PPS)

- **Traditional PS**
  - Dabigatran ~ age + gender + DM + HTN + NSAID use….
  - Balance achieved on all included covariates; so if all individual risk factors for stroke and bleeding are added to this model, adjusted estimates for dabigatran vs warfarin for both outcomes can be calculated from a single PS matched cohort
- **Prognostic PS**
  - Dabigatran ~ $DRS_{Bleeding}$ + DRS_{Stroke}
  - Balances disease risks for both the outcomes; which in turn will induce 'prognostic balance' with respect to risk factors for both outcomes
  - Similar to standard PS, adjusted estimates for dabigatran vs warfarin for both outcomes can be calculated from a single PS matched cohort

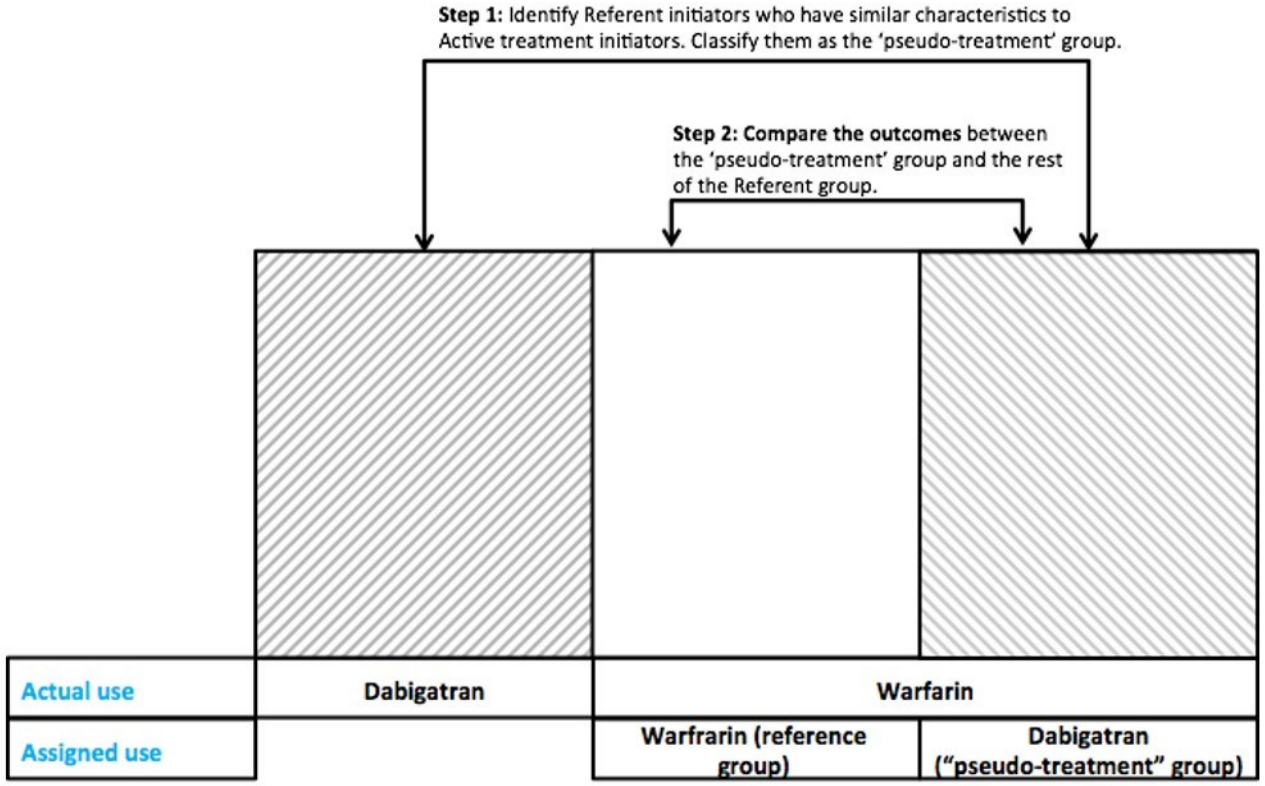
Approach 2: Mahalanobis-distance matching

- A widely used measure for generalized distance between two points in a multivariate space
- For matching, M-distance calculated between individual treated and control units
- $\Delta^2 = (X - m)\, C^{-1}\, (X-m)^T$
  - X: row vector consisting of the multivariate measurement (ie. DRS for stroke and DRS for bleeding) for a treated observation
  - m: row vector consisting of the multivariate measurement (ie. DRS for stroke and DRS for bleeding) for a reference observation
  - C^{-1}: Inverse covariance matrix of variables
  - $(X-m)^T$: Transpose of the matrix (X-m)
- Based on this metric, each treated unit is matched to the closest control unit without resampling

Approach 2: Mahalanobis-distance matching



Evaluating performance of both methods with dry run



Any association between the assigned treatment and outcome in the pseudo-population is due to confounding; success of a DRS model is evaluated based on its ability to retrieve unconfounded null results in the dry run analysis.

Wyss et al. *AJE* 2017

Evaluating performance of both methods in sequential monitoring

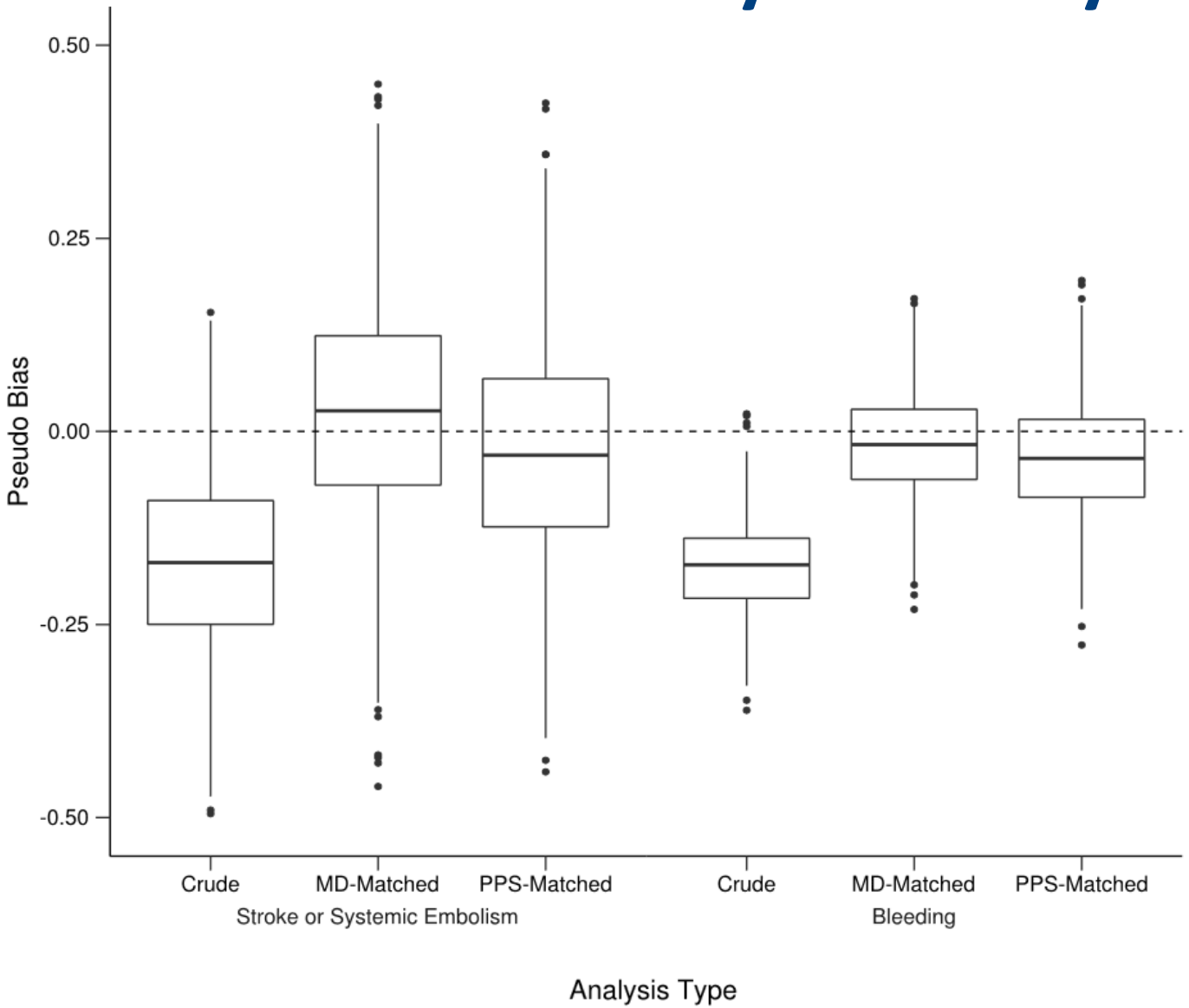
- Beginning on the date of dabigatran market entry, conducted sequential monitoring for 90 day periods for 12 total periods
- In each period, we identified initiators of dabigatran and warfarin; conducted 1:1 matching on the PPS and M-distance. As a comparison, also conducted 1:1 matching on traditional PS
  - Traditional PS-model: 72 covariates
  - PPS model: 2 covariates, their squared and interaction terms
- For each subsequent monitoring period, additional follow-up data for outcome assessment were added to all pairs matched in earlier periods and new matched pairs were pooled with previously matched pairs to conduct a cumulative analysis
  - Outcomes- Stroke and bleeding
  - Follow-up approach- As treated
  - Analysis: Cox PH models

Results

Patient characteristics

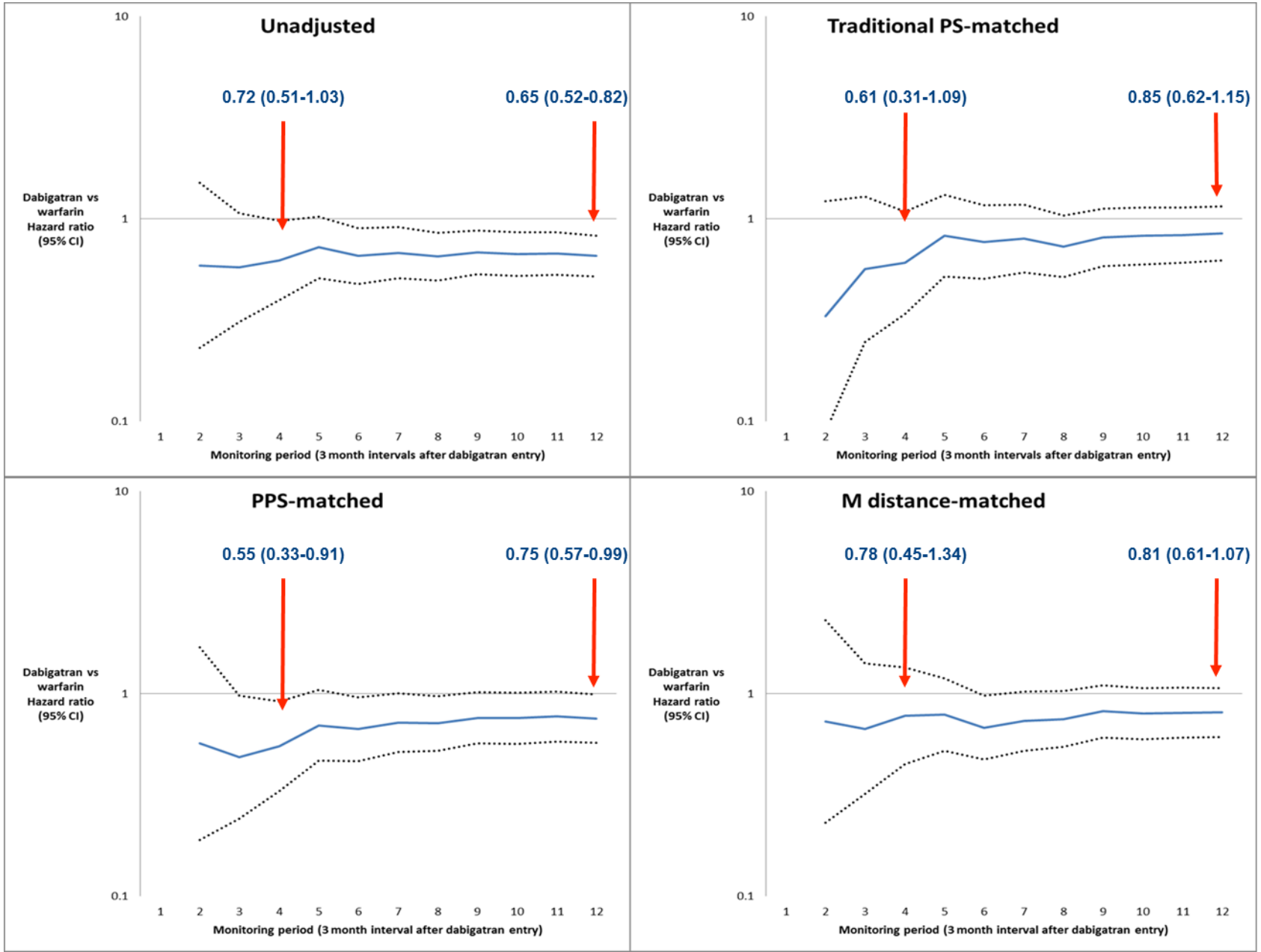
	Warfarin	Dabigatran	Standardized difference
Patients (n)	(56,456)	(22,809)	
Age (Mean (±SD))	71.10 (±12.13)	67.29 (±12.23)	-0.31
Female	22229(39.37)	8209(35.99)	-0.07
Comorbid conditions			
Systemic embolism	728(1.29)	112(0.49)	-0.09
Deep vein thrombosis	4241(7.51)	289(1.27)	-0.31
Pulmonary embolism	2932(5.19)	103(0.45)	-0.29
Hypertension	53873(95.42)	22061(96.72)	0.07
Hyperlipidemia	26638(47.18)	10628(46.60)	-0.01
Heart failure	12464(22.08)	3648(15.99)	-0.16
Ischemic stroke	5144(9.11)	1599(7.01)	-0.08
Transient ischemic attack	2637(4.67)	947(4.15)	-0.03
Myocardial infarction	3180(5.63)	874(3.83)	-0.08
Peripheral vascular disease	2675(4.74)	665(2.92)	-0.10
Diabetes	14242(25.23)	4774(20.93)	-0.10
Intracranial bleeding	125(0.22)	35(0.15)	-0.02
Peptic ulcer disease	8414(14.90)	3091(13.55)	-0.04
Lower/unspecified gastrointestinal bleed	2036(3.61)	514(2.25)	-0.08
Upper gastrointestinal bleed	345(0.61)	62(0.27)	-0.05
Urogenital bleed	21(0.04)	10(0.04)	0.00
Other bleeds	2538(4.50)	414(1.82)	-0.15
Atherosclerosis	17729(31.40)	5964(26.15)	-0.12

Results from the dry run analysis

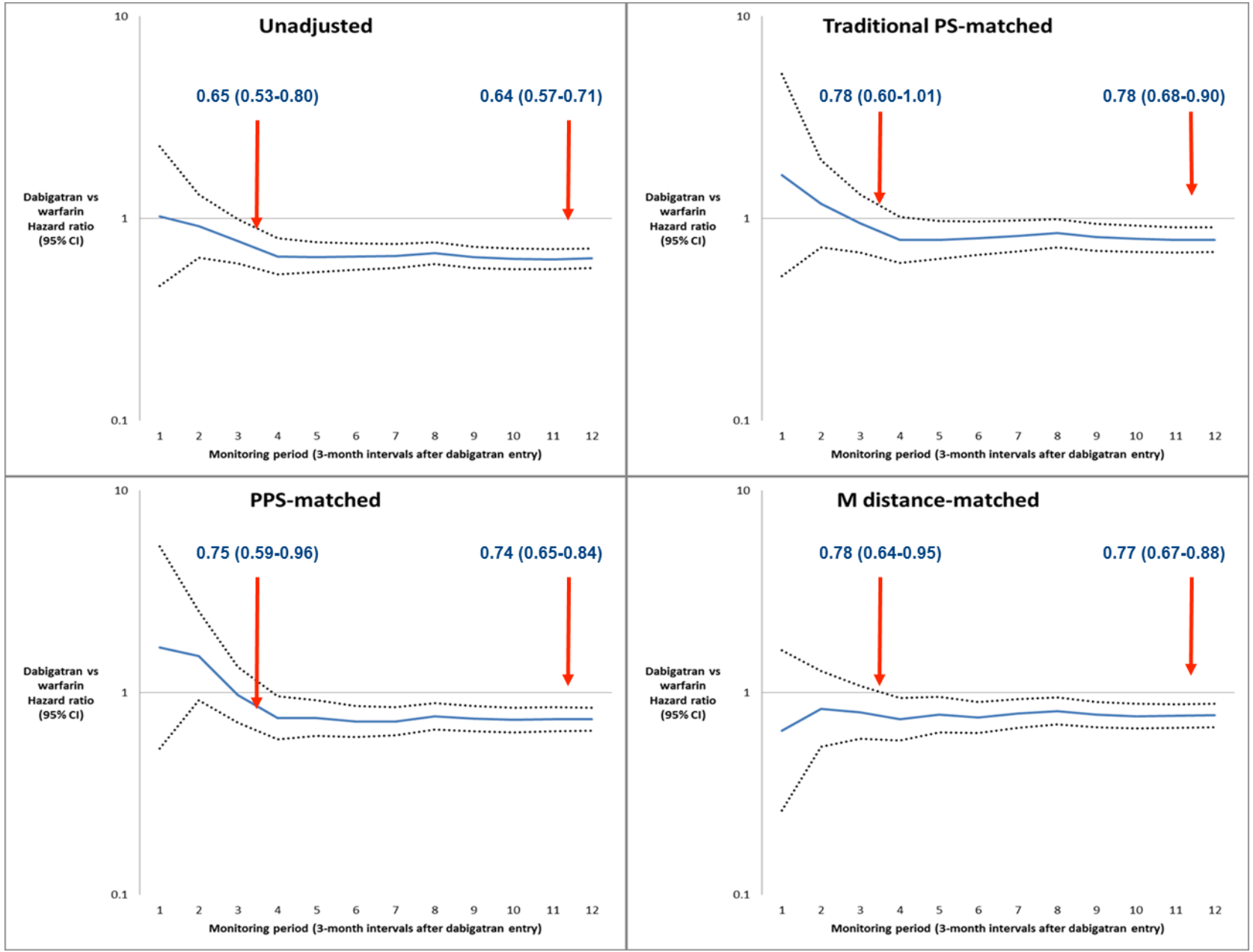


Summary: PPS matching on average resulted in greater residual confounding for both outcomes compared to M-distance matching

Results for sequential monitoring of the stroke outcome



Results for sequential monitoring of the bleeding outcome



Conclusions

- Combining two DRSs using a prognostic PS or M-distance matching allowed to adjust for confounding in analysis of both the outcomes
  - M-distance matching produced estimates that were close to traditional PS-matching, but PPS-matching resulted in estimates with potential residual confounding compared to traditional PS-matching
  - Further, M-distance matching on multiple DRS appeared to produce more stable results in the very early marketing period compared to both PPS and traditional PS matching, as demonstrated by higher precision and proximity to the estimates from later periods, suggesting this approach maybe preferable
- Limitations of DRS approach include limited utility for rare outcomes, especially the ones where risk factors are not well-understood

Thank you

Study team: Rishi J. Desai, Richard Wyss, Yinzhu Jin, Justin Bohn, Sengwee Toh, Austin Cosgrove, Adee Kennedy, Jessica Kim, Clara Kim, Rita Ouellet-Hellstrom, Sara Karami, Jacqueline M. Major, Aaron Niman, Shirley V. Wang, and Joshua J. Gagne

rdesai@bwh.harvard.edu @Rishidesai11