OHDSI Methods for Causal Effect Estimation

George Hripcsak, David Madigan, Patrick Ryan, Martijn Schuemie, Marc Suchard

http://www.ohdsi.org

"The sole cause and root of almost every defect in the sciences is this: that whilst we falsely admire and extol the powers of the human mind, we do not search for its real helps."

— Novum Organum: Aphorisms [Book One], 1620, Sir Francis Bacon

Thromb Haemost 2016; 116(05): 975-986
DOI: 10.1160/TH16-05-0403

Stroke, Systemic or Venous Thromboembolism

Schattauer GmbH

Real-world comparison of major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran, rivaroxaban, or warfarin

A propensity score matched analysis

Gregory Y. H. Lip, Allison Keshishian, Shital Kamble, Xianying Pan, Jack Mardekian, Ruslan Horblyuk, Melissa Hamilton

When comparisons were made between NOACs, matched rivaroxaban patients had a significantly higher risk of major bleeding (HR: 1.82; 95 % CI: 1.36–2.43) compared to apixaban patients.

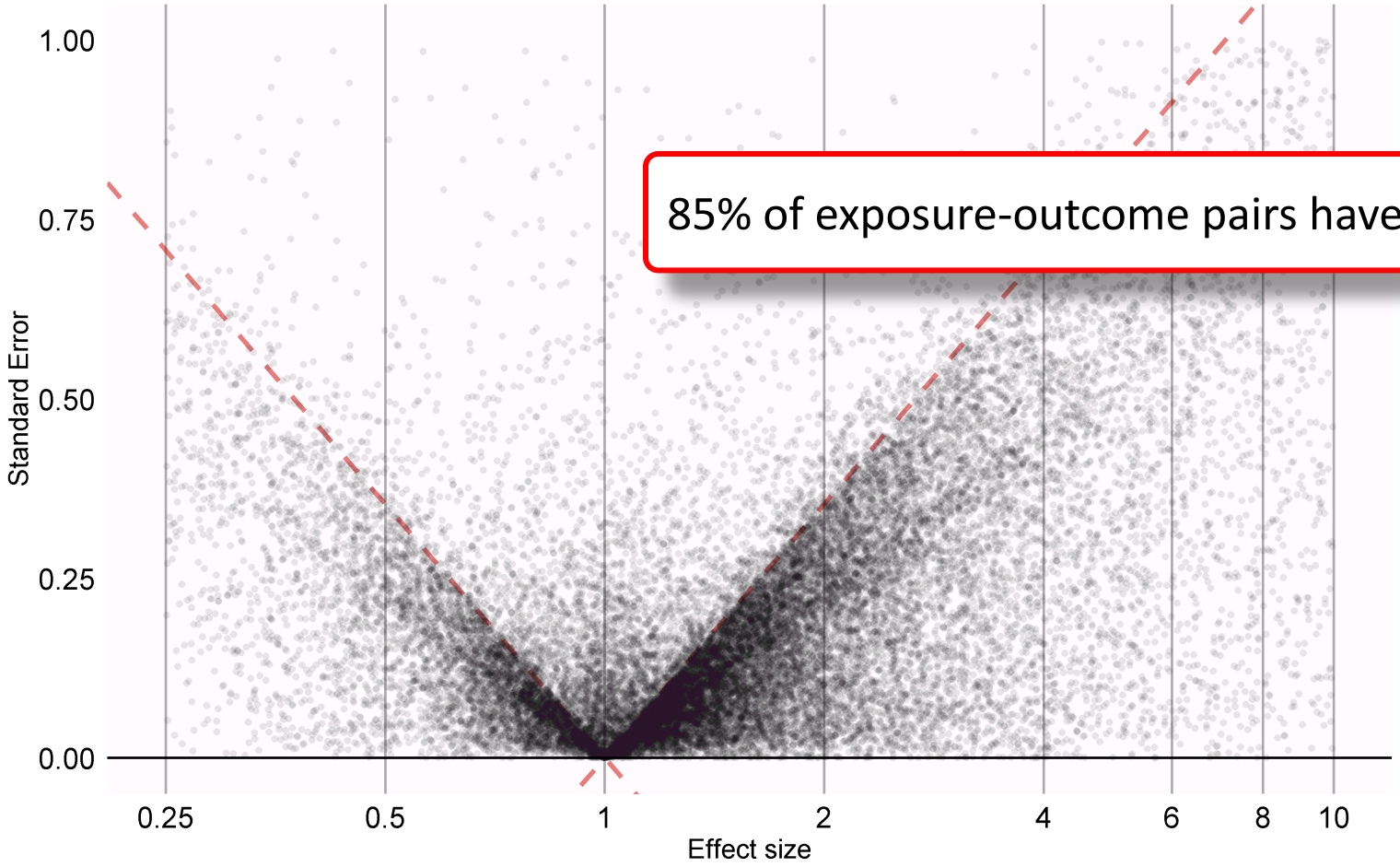
Reliability: Analysis Ignores…

- Selection bias
- Measurement error
- Model misspecification
- Multiple modeling
- Unmeasured confounding

"Grave errors are commonplace, perhaps typical. It does no good to append a claim that you have included in the regression all relevant covariates, a claim that there are no unmeasured confounders and that you could not be mistaken in making this claim. Who are you that you could not be mistaken?"

- Paul Rosenbaum

Observational research results in literature



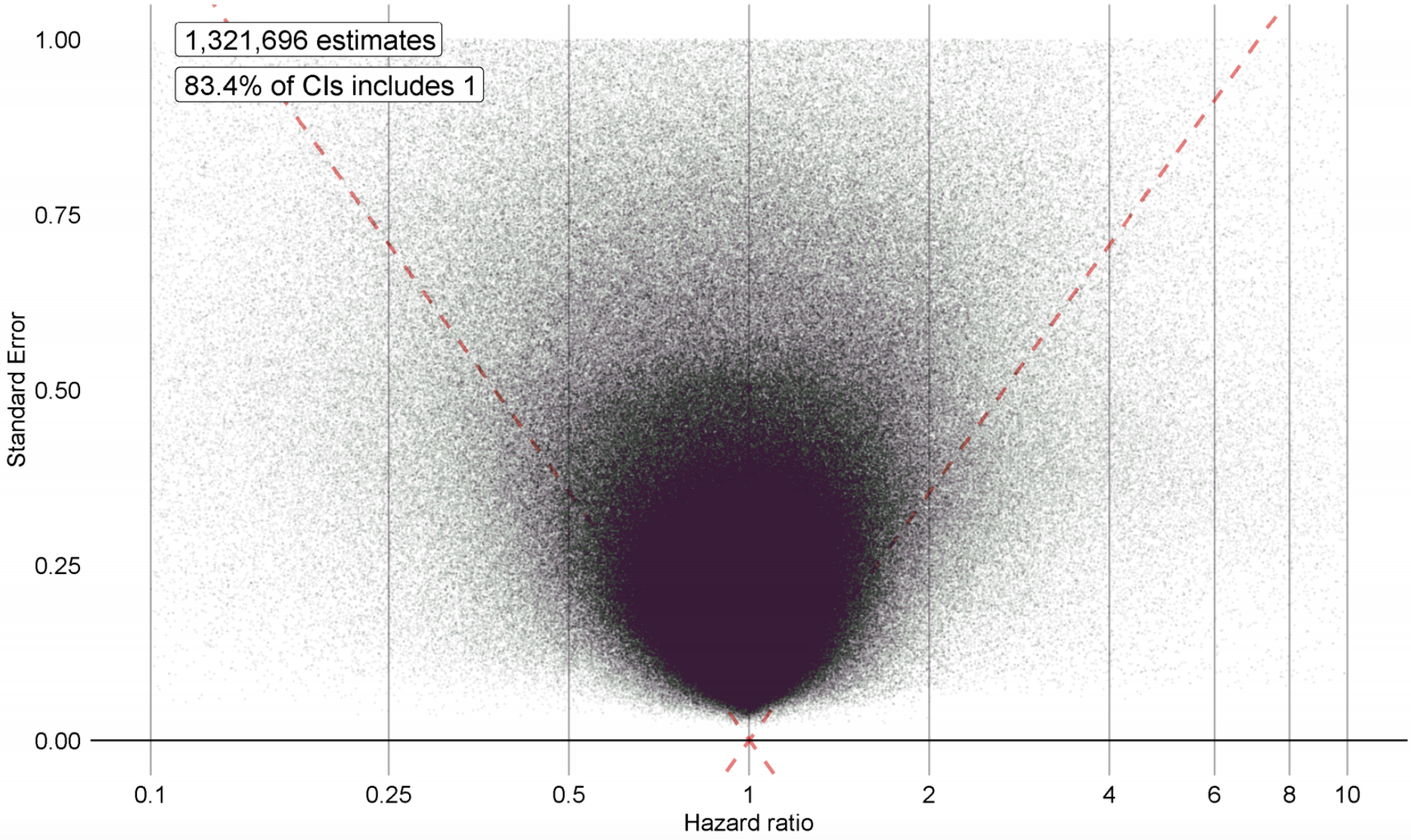
85% of exposure-outcome pairs have $p < 0.05$

29,982 estimates
11,758 papers

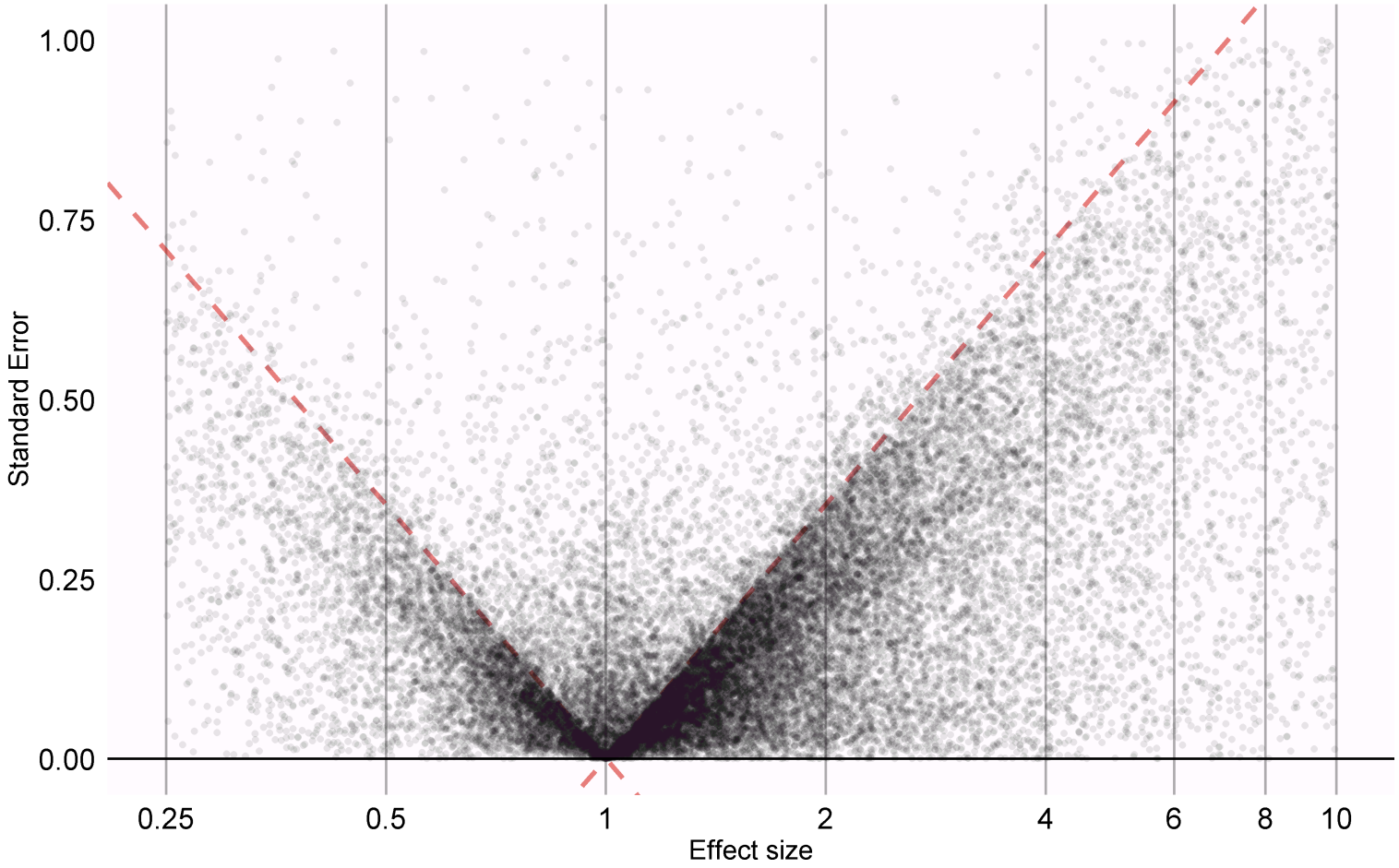
A New Approach

- Reproducible, systematized, open source approach at scale
- Negative controls
  - Drugs and outcomes “known” to have no causal association
  - Literature, product labels, spontaneous reports
  - Empirical p-values
- Positive Controls
  - Inject signals onto negative controls with known effect size
  - Calibrated confidence intervals

LEGEND results



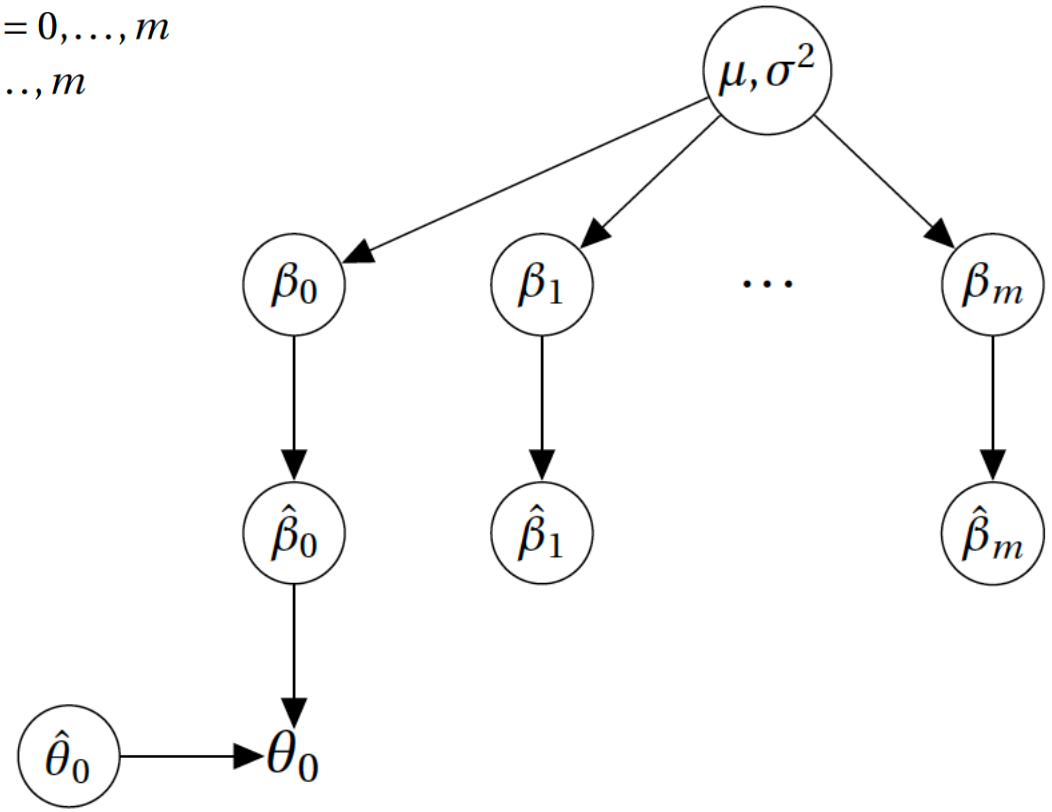
Observational research results in literature



29,982 estimates
11,758 papers

Calibration Assumptions

θ_0	true effect size of interest
θ_i	true effect size for the controls, $i = 1, \ldots, m$
$\hat{\theta}_i$	estimated effect sizes, $i = 0, \ldots, m$
$\hat{\beta}_i = \hat{\theta}_i - \theta_i$	"estimated bias," $i = 0, \ldots, m$
$E[\hat{\beta}_i] = \beta_i$	$i = 0, \ldots, m$



Many models, many databases

$$p(\theta_0|\mathscr{D}) = \sum_k p(\theta_0|M_k,\mathscr{D})\,p(M_k|\mathscr{D})$$

where the data, $\mathscr{D}$, comprise:

$$\hat{\theta}^k_{0j}, k = 1,\ldots,M, j = 1,\ldots,D$$

$$\hat{\theta}^k_{ij}, k = 1,\ldots,M, j = 1,\ldots,D, i = 1,\ldots,Q$$

$$\theta_i, i = 1,\ldots,Q$$

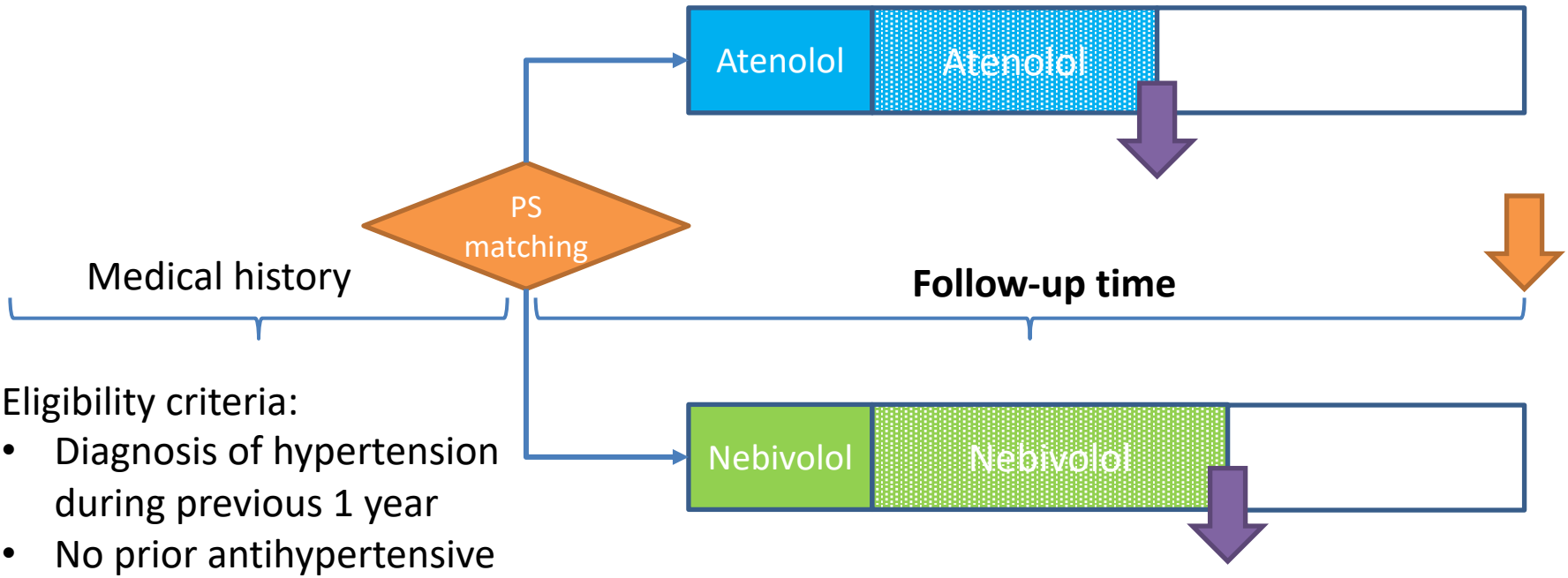
can show that: $p(M_k|\mathscr{D}) \propto \prod_{i=1}^{Q}\prod_{j=1}^{D} p(\hat{\theta}^k_{ij}|\theta_i, M_k)$

Combining calibration with random effects meta-analysis and BMA

Method: Study design (LEGEND)

Treatment strategies:
- Atenolol
- Nebivolol

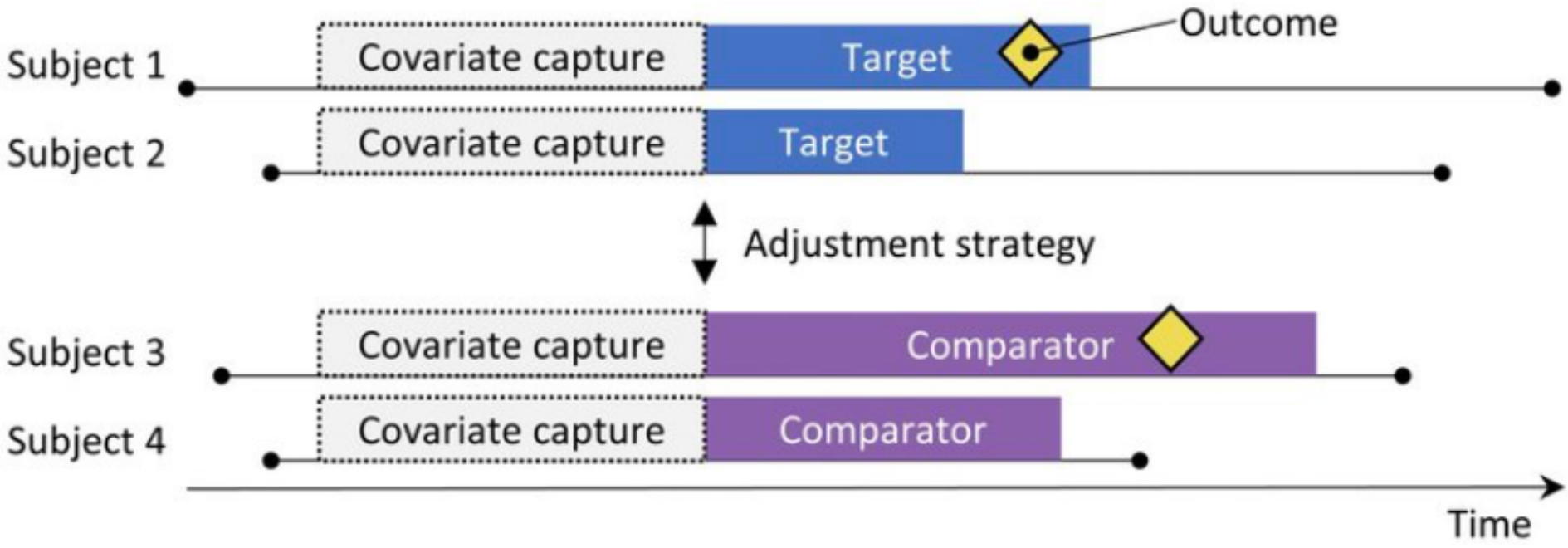
Causal contrasts of interest:
- On-treatment effect
- Intent-to-treat effect



Eligibility criteria:
- Diagnosis of hypertension during previous 1 year
- No prior antihypertensive drug
- No prior cardiovascular outcome

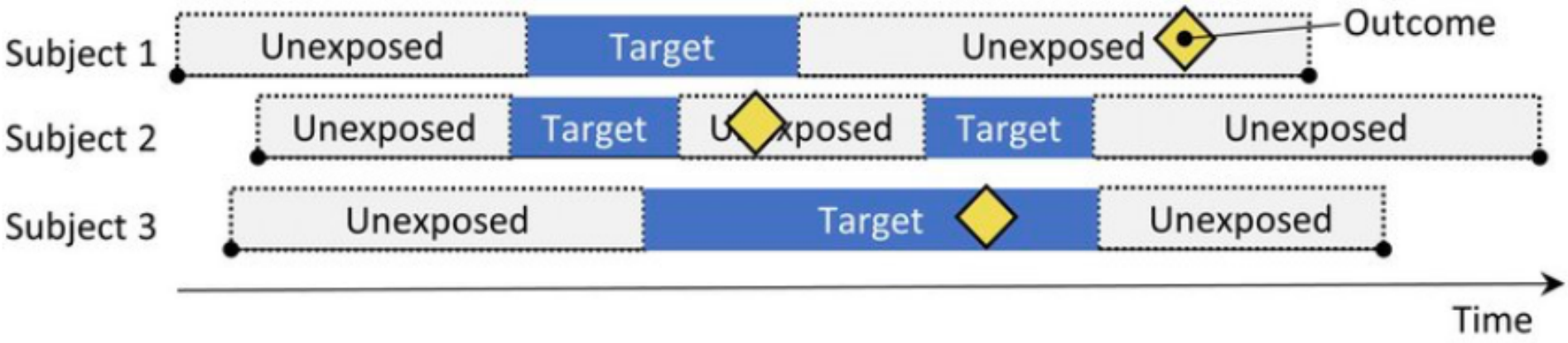
Outcome (Major Adverse Cardio-Cerebrovascular Event):
- Hospitalized myocardial infarction, heart failure, stroke and sudden cardiac death

Cohort Methods



Other Methods: Case Control

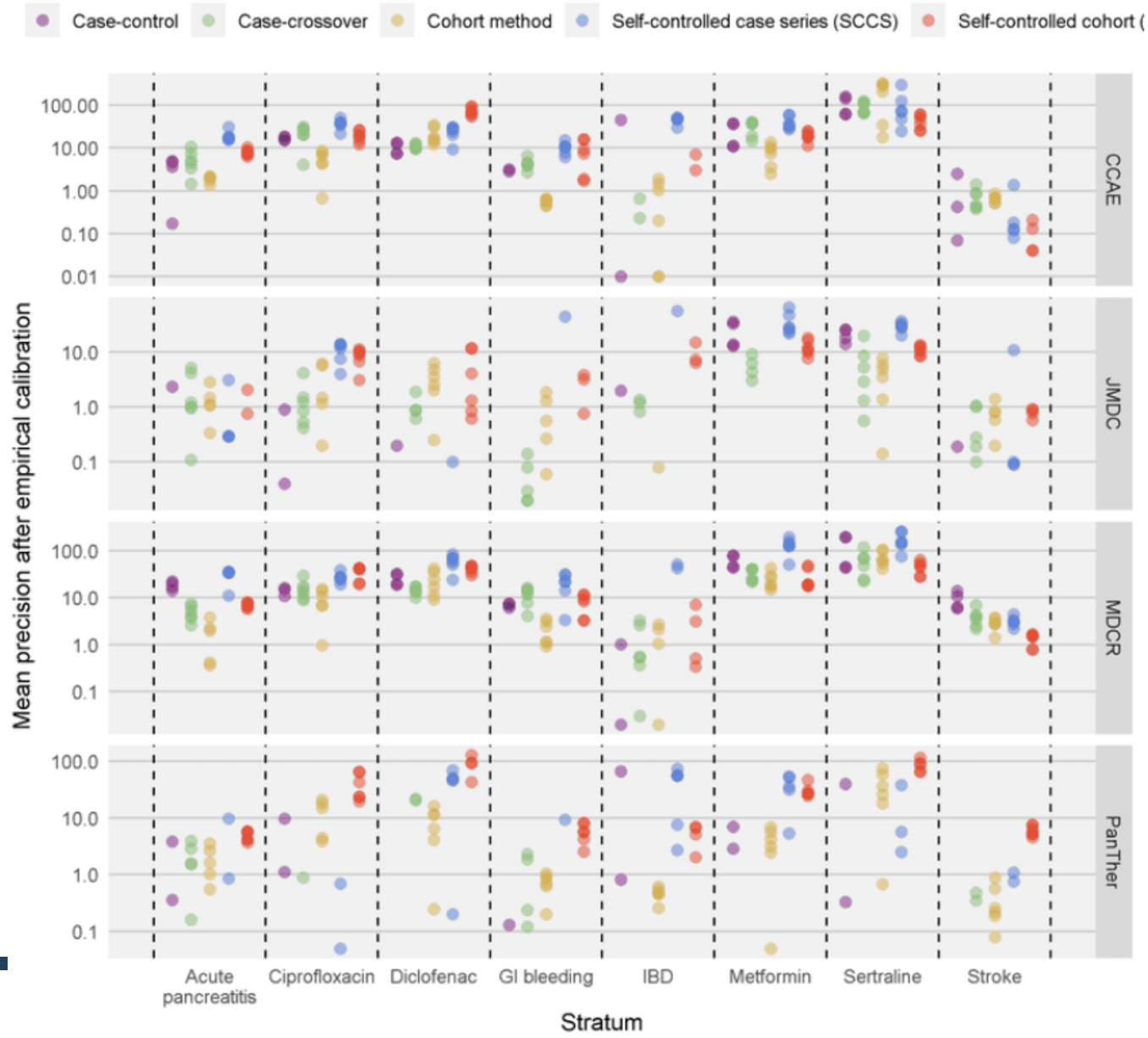
Other Methods: SCCS



Gold Standard Performance: CCAE

Analysis choices	AUC	95% CI coverage	Mean precision	MSE	Type 1 error	Type 2 error	Non-estimable
Cohort method							
No PS, simple outcome model	0.78	0.92	3.35	0.51	0.08	0.60	0.19
1-on-1 matching, unstratified outcome model	0.78	0.89	9.35	0.33	0.08	0.37	0.31
Variable ratio matching, stratified outcome model	0.78	0.92	7.45	0.34	0.06	0.41	0.30
Stratification	0.79	0.90	8.68	0.41	0.10	0.35	0.26
IPTW	0.78	0.93	4.14	0.44	0.08	0.55	0.24
Var ratio matching + full outcome model	0.78	0.93	7.76	0.31	0.04	0.43	0.45
Self-controlled cohort (SCC)							
Time exposed, inc. exp. start date	0.87	0.92	12.28	0.49	0.09	0.29	0.07
30 days, incl. exp. start date	0.87	0.92	12.25	0.19	0.09	0.34	0.04
Time exposed, inc. exp. start date, require full obs.	0.88	0.92	13.91	0.47	0.09	0.27	0.08
Time exposed, ex. exp. start date	0.87	0.94	11.75	0.22	0.08	0.18	0.23
30 days, ex. exp. start date	0.89	0.93	14.46	0.14	0.08	0.29	0.05
Time exposed, ex. exp. start date, require full obs.	0.87	0.94	13.04	0.20	0.07	0.17	0.24
Case-control							
2 controls per case	0.84	0.92	7.35	0.59	0.08	0.48	0.04
10 controls per case	0.84	0.92	7.37	0.62	0.08	0.48	0.01
Nesting in indication, 2 controls per case	0.87	0.91	12.90	0.54	0.10	0.34	0.01
Nesting in indication, 10 controls per case	0.86	0.92	12.28	0.55	0.10	0.35	0.02
Case-crossover							
Simple case-crossover, -30 days	0.87	0.92	10.17	0.46	0.08	0.43	0.04
Simple case-crossover, -180 days	0.85	0.93	10.79	0.61	0.07	0.44	0.02
Nested case-crossover, -30 days	0.87	0.92	11.78	0.41	0.08	0.34	0.06
Nested case-crossover, -180 days	0.86	0.93	12.26	0.55	0.06	0.36	0.03
Nested case-time-control, -30 days	0.87	0.92	10.76	0.38	0.06	0.38	0.05
Nested case-time-control, -180 days	0.87	0.94	10.80	0.21	0.07	0.35	0.19
Self-controlled case series (SCCS)							
Simple SCCS	0.90	0.95	15.78	0.17	0.07	0.20	0.20
Including day 0	0.87	0.93	11.67	0.53	0.08	0.39	0.09
Using pre-exposure window	0.90	0.95	12.98	0.19	0.08	0.24	0.20
Using age and season	0.91	0.94	22.15	0.16	0.08	0.19	0.20
Using event-dependent observation	0.88	0.95	12.12	0.18	0.08	0.24	0.20
Using all other exposures	0.91	0.95	21.98	0.16	0.05	0.11	0.20

Gold Standard Performance: CCAE



Concluding thoughts

- An international community and global data network can be used to generate real-world evidence in a secure, reliable and efficient manner
- Common data model critically important
- Much work remains on establishing (and improving) actual operating characteristics of current approaches to causal inference