

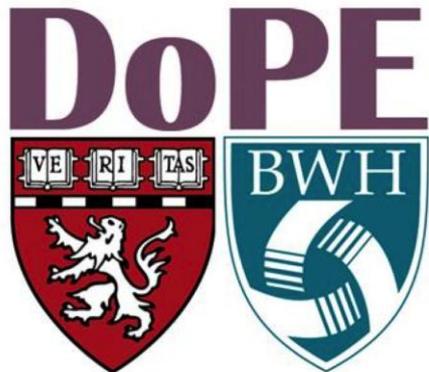
Integrating Effectiveness and Safety Outcomes in the Assessment of Treatments

Jessica M. Franklin

Instructor in Medicine

Division of Pharmacoepidemiology & Pharmacoeconomics

Brigham and Women's Hospital and Harvard Medical School



Making decisions about treatments

- At the policy level:
 - FDA market approval for drugs and devices
 - FDA's Sentinel Initiative for post-approval safety surveillance
 - Payer reimbursement (CMS, medical/pharmacy insurance)

Making decisions about treatments

- At the policy level:
 - FDA market approval for drugs and devices
 - FDA's Sentinel Initiative for post-approval safety surveillance
 - Payer reimbursement (CMS, medical/pharmacy insurance)
- At the patient level (PCORI's mandate):
 - "Given my personal characteristics, conditions and preferences, what should I expect will happen to me?"
 - "What are my options and what are the potential benefits and harms of those options?"
 - "What can I do to improve the outcomes that are most important to me?"
 - "How can clinicians and the care delivery systems they work in help me make the best decisions about my health and healthcare?"

Example question

- Randomized trials show that a new drug:
 - Reduces the risk of heart attack by approximately 15%
 - Increases the risk of a painful adverse event by 10x.
- Should the drug be approved?
- As a patient, should I take this drug?

Example question

- Randomized trials show that a new drug:
 - Reduces the risk of heart attack by approximately 15%
 - Increases the risk of a painful adverse event by 10x.
- Should the drug be approved?
- As a patient, should I take this drug?

→ This is the context for most statins.

- The adverse event (rhabdomyolysis) is extremely rare even among patient taking a statin ($\sim 0.44 / 10,000$ patients).
- Probably worth it for a patient at increased risk of MI.
- Probably not a good idea for a patient with low risk of MI.

Example question

- Randomized trials show that a new drug:
 - Reduces the risk of heart attack by approximately 15%
 - Increases the risk of a painful adverse event by 10x.
- Should the drug be approved?
- As a patient, should I take this drug?

→ This is the context for most statins.

- The adverse event (rhabdomyolysis) is extremely rare even among patient taking a statin ($\sim 0.44 / 10,000$ patients).
- Probably worth it for a patient at increased risk of MI.
- Probably not a good idea for a patient with low risk of MI.
- Cerivastatin was found to have higher risk of rhabdo (16—80 times higher than other statins) and was withdrawn in 2001.

Observational studies of safety and effectiveness

- Nonrandomized studies are often necessary for evaluating multiple outcomes:
 - Trials are usually underpowered to detect differences in relatively rare safety outcomes
 - Trials generally enroll healthier patient populations with lower overall risk of safety events or drug interactions
 - Effectiveness of drug may differ in general and trial populations
- Nonrandomized studies come with additional challenges:
 - Treatments must be compared in patients that are balanced with respect to their baseline risk of all outcomes of interest.
 - A covariate that strongly influences one outcome may be an instrument for another outcome – so balance on that covariate can decrease bias in one analysis while increasing it in another.

Example: Equivalence trials

- Equivalent efficacy across statins have been established in several RCTs and meta-analyses¹:
 - Reduces LDL-C $\geq 40\%$: rosuvastatin 10 mg or higher and atorvastatin 20 mg or higher
 - Reduces LDL-C 30—40%: atorvastatin 10 mg, fluvastatin 80 mg, lovastatin 40/80 mg, and simvastatin 20 mg
 - Reduces LDL-C 20—30%: fluvastatin 40 mg, lovastatin 10/20 mg, pravastatin 20/40 mg, and simvastatin 10 mg
- How does safety differ across statins?
 - Trials were underpowered to detect.
 - Observational studies can evaluate safety across statin doses with similar effectiveness to aid in treatment decisions.
 - If equivalent effectiveness is known, patients initiating different but equivalent treatments may be similar, leading to valid comparisons.
 - If safety differences are known, patients may be similar in terms of cardiovascular risk, but different with respect to safety profile.

1. Weng et al., Journal of Clinical Pharmacy and Therapeutics. 2010; 35(2): 139–151.

Example: Composite endpoints

- Comparative safety of analgesics in older adults with arthritis¹
- Medicare beneficiaries in PA and NJ were identified when they initiated an analgesic for treating pain:
 - Nonselective NSAIDs
 - Coxibs
 - Opioids
- Safety outcomes of interest included:
 - Cardiovascular events (MI, stroke, heart failure, revascularization, cardiac death)
 - Gastrointestinal events (GI bleed, bowel obstruction)
 - Kidney injury (renal failure, dialysis)
 - Liver injury
 - Fractures
 - Falls

1. .Solomon et al., Archives of Internal Medicine, 2010.

Example: Composite endpoints

- Patients were matched 1:1:1 on the basis of 2 propensity scores:
 - Estimated probability of Coxib use vs ns-NSAID use.
 - Estimated probability of opioid use vs ns-NSAID use.
- PS models included the same list of confounders measured in claims data in the 365 days prior to exposure:
 - Demographics
 - Cardiovascular diagnoses and medication use
 - Osteoporosis and fracture diagnoses and medication use
 - Gastrointestinal diagnoses and treatments
 - Liver or renal disease diagnoses

Example: Results

Table 3. Safety Events Among Propensity Score–Matched Older Adults With Arthritis Initiating Prescription Analgesic Treatment

Adverse Event ^a	HR (95% CI)		
	nsNSAIDs	Coxibs	Opioids
		Individual Safety Events	
Myocardial infarction	1 [Reference]	1.63 (0.96-2.77)	2.25 (1.32-3.84)
Heart failure	1 [Reference]	1.26 (0.87-1.84)	1.63 (1.12-2.38)
Stroke	1 [Reference]	1.04 (0.66-1.63)	0.91 (0.55-1.50)
Coronary revascularization	1 [Reference]	3.30 (1.47-7.38)	5.34 (2.40-11.90)
Out-of-hospital cardiac death	1 [Reference]	0.86 (0.44-1.71)	1.96 (1.05-3.67)
Upper GI tract bleeding	1 [Reference]	0.52 (0.27-1.02)	1.03 (0.55-1.91)
Lower GI tract bleeding	1 [Reference]	0.61 (0.27-1.36)	0.95 (0.43-2.09)
Bowel obstruction	1 [Reference]	1.89 (0.50-7.13)	4.87 (1.40-17.02)
Acute kidney injury	1 [Reference]	0.98 (0.71-1.35)	1.53 (1.12-2.09)
Falls	1 [Reference]	0.73 (0.47-1.14)	1.64 (1.09-2.47)
Fracture			
Hip	1 [Reference]	1.35 (0.51-3.57)	3.02 (1.20-7.58)
Humerus	1 [Reference]	1.06 (0.39-2.85)	9.26 (4.25-20.18)
Pelvic	1 [Reference]	1.68 (0.73-3.86)	3.13 (1.42-6.91)
Radius	1 [Reference]	0.50 (0.22-1.14)	3.74 (2.12-6.60)
Hepatotoxic effects	1 [Reference]	1.08 (0.58-2.01)	1.14 (0.58-2.22)

Example: Results

Table 3. Safety Events Among Propensity Score–Matched Older Adults With Arthritis Initiating Prescription Analgesic Treatment

Adverse Event ^a	HR (95% CI)		
	nsNSAIDs	Coxibs	Opioids
		Individual Safety Events	
Myocardial infarction	1 [Reference]	1.63 (0.96-2.77)	2.25 (1.32-3.84)
Heart failure	1 [Reference]	1.26 (0.87-1.84)	1.63 (1.12-2.38)
Stroke	1 [Reference]	1.04 (0.66-1.63)	0.91 (0.55-1.50)
Coronary revascularization	1 [Reference]	3.30 (1.47-7.38)	5.34 (2.40-11.90)
Out-of-hospital cardiac death	1 [Reference]	0.86 (0.44-1.71)	1.96 (1.05-3.67)
Upper GI tract bleeding	1 [Reference]	0.52 (0.27-1.02)	1.03 (0.55-1.91)
Lower GI tract bleeding	1 [Reference]	0.61 (0.27-1.36)	0.95 (0.43-2.09)
Bowel obstruction	1 [Reference]	1.89 (0.50-7.13)	4.87 (1.40-17.02)
Acute kidney injury	1 [Reference]	0.98 (0.71-1.35)	1.53 (1.12-2.09)
Falls	1 [Reference]	0.73 (0.47-1.14)	1.64 (1.09-2.47)
Fracture			
Hip	1 [Reference]	1.35 (0.51-3.57)	3.02 (1.20-7.58)
Humerus	1 [Reference]	1.06 (0.39-2.85)	9.26 (4.25-20.18)
Pelvic	1 [Reference]	1.68 (0.73-3.86)	3.13 (1.42-6.91)
Radius	1 [Reference]	0.50 (0.22-1.14)	3.74 (2.12-6.60)
Hepatotoxic effects	1 [Reference]	1.08 (0.58-2.01)	1.14 (0.58-2.22)

Example: Results

Outcomes were combined to create 6 composite endpoints.

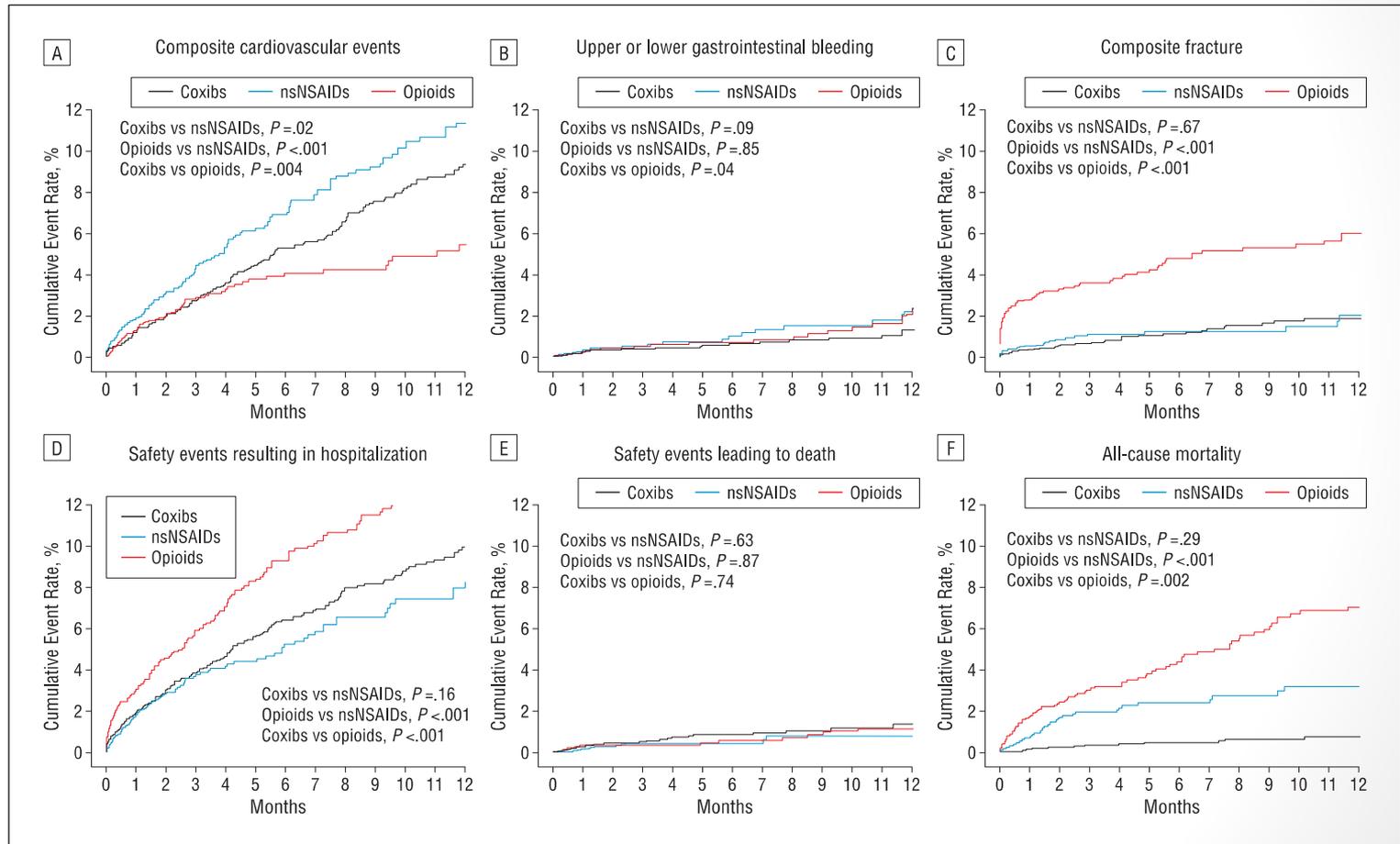


Figure. Kaplan-Meier curves for the cumulative incidence of the 6 composite safety measures. A, Composite cardiovascular events. B, Upper or lower gastrointestinal tract bleeding. C, Composite fracture. D, Any of the individual safety events resulting in hospitalization. E, Any of the individual safety events leading to immediate death or a hospitalization with death. F, All-cause mortality. P values were determined with the log-rank test. Coxibs indicates selective cyclooxygenase-2 inhibitors; nsNSAIDs, nonselective nonsteroidal anti-inflammatory drugs.

Example: Results

Outcomes were combined to create 6 composite endpoints.

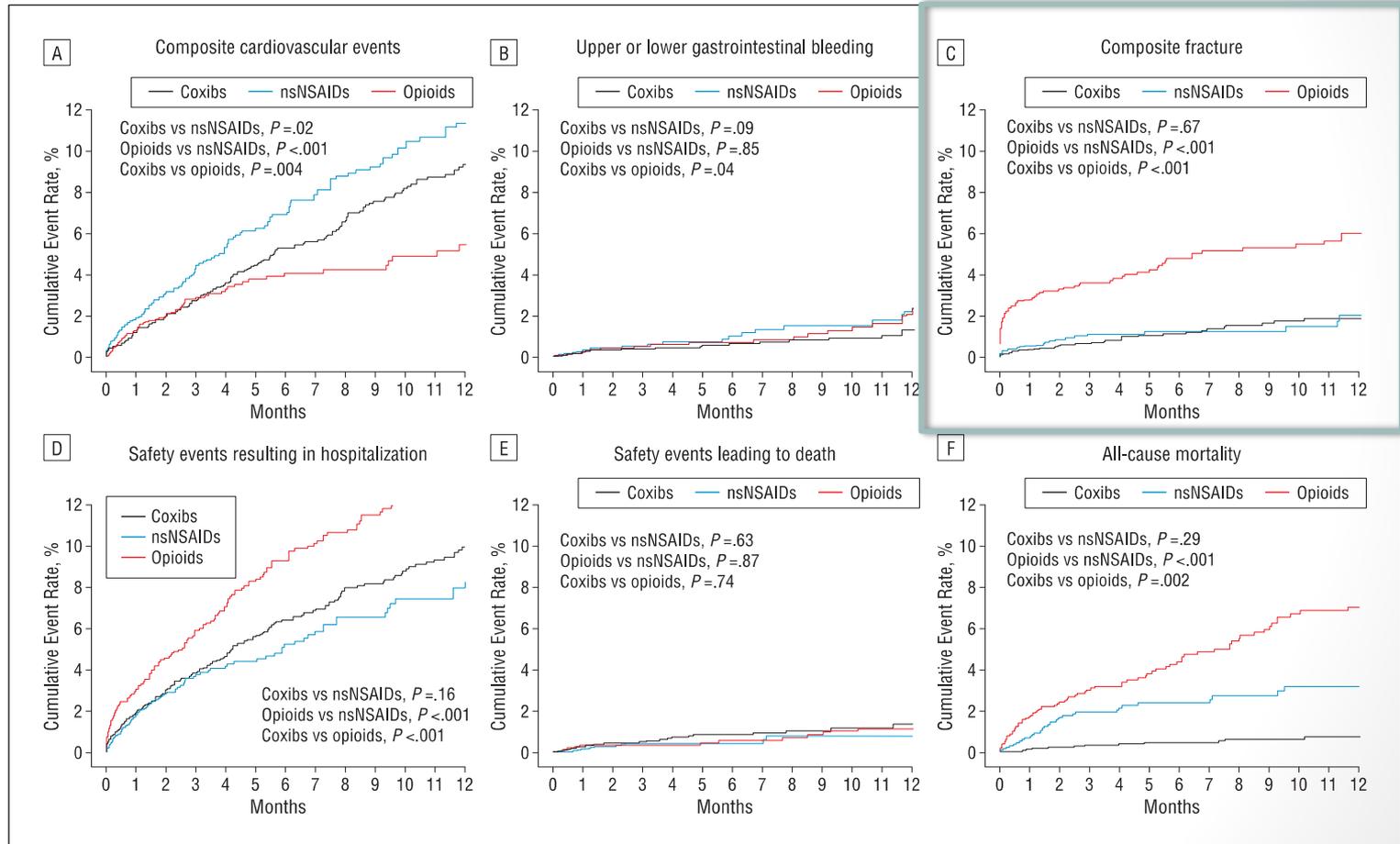


Figure. Kaplan-Meier curves for the cumulative incidence of the 6 composite safety measures. A, Composite cardiovascular events. B, Upper or lower gastrointestinal tract bleeding. C, Composite fracture. D, Any of the individual safety events resulting in hospitalization. E, Any of the individual safety events leading to immediate death or a hospitalization with death. F, All-cause mortality. P values were determined with the log-rank test. Coxibs indicates selective cyclooxygenase-2 inhibitors; nsNSAIDs, nonselective nonsteroidal anti-inflammatory drugs.

Example: Results

Outcomes were combined to create 6 composite endpoints.

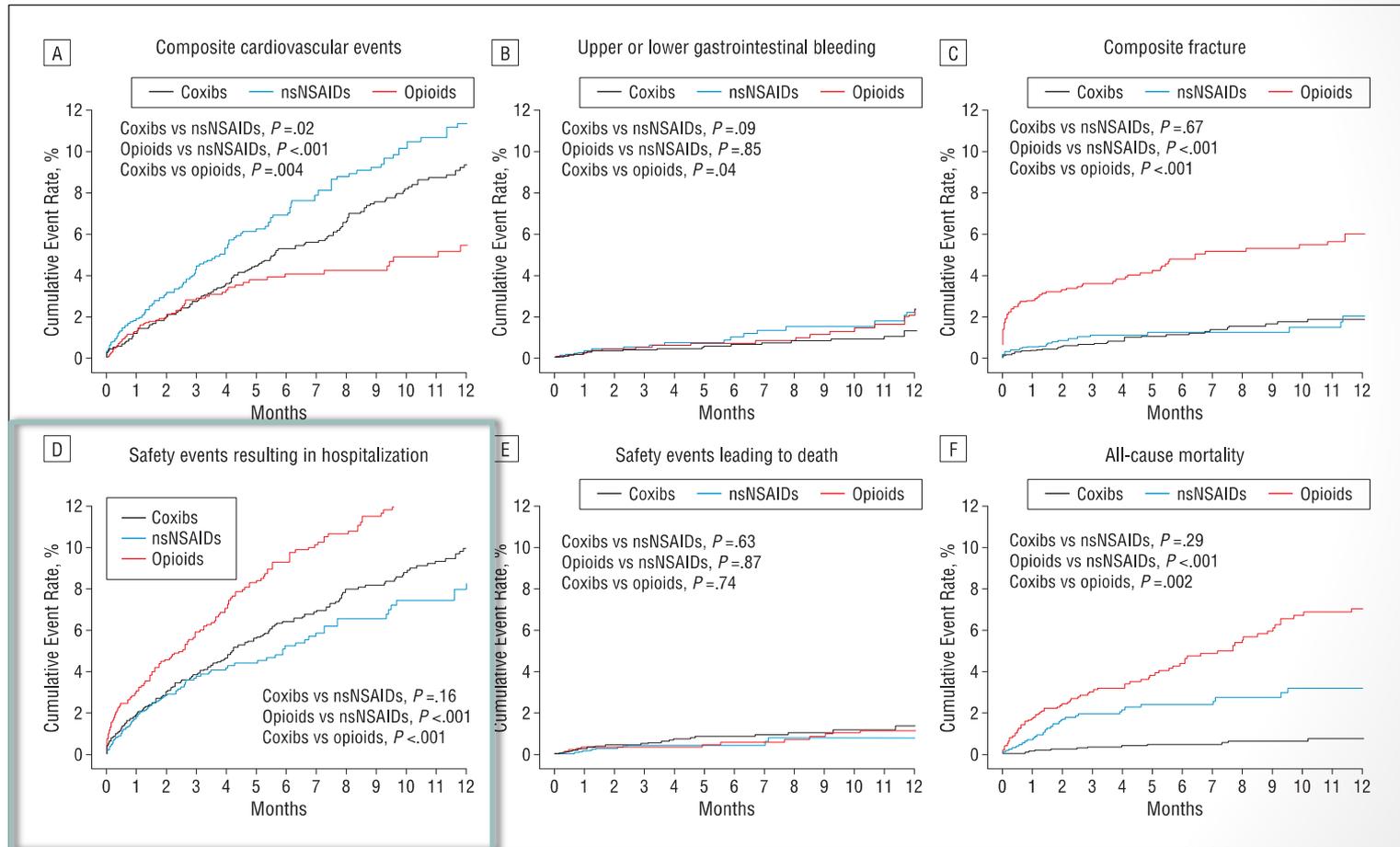


Figure. Kaplan-Meier curves for the cumulative incidence of the 6 composite safety measures. A, Composite cardiovascular events. B, Upper or lower gastrointestinal tract bleeding. C, Composite fracture. D, Any of the individual safety events resulting in hospitalization. E, Any of the individual safety events leading to immediate death or a hospitalization with death. F, All-cause mortality. P values were determined with the log-rank test. Coxibs indicates selective cyclooxygenase-2 inhibitors; nsNSAIDs, nonselective nonsteroidal anti-inflammatory drugs.

Composite endpoints

- Works for combining multiple safety outcomes, but still doesn't integrate effectiveness.
- Assumes all events leading to hospitalization are of equal importance
 - Equates hospital stays of 1 day and 1 year.
- Doesn't provide a decision:
 - If one drug is 2% worse, are they equivalent from a policy standpoint?

Making decisions about treatments

- At the policy level:
 - FDA market approval for drugs and devices
 - FDA's Sentinel Initiative for post-approval safety surveillance
 - Payer reimbursement (CMS, medical/pharmacy insurance)
- At the patient level (PCORI's mandate):
 - "Given my personal characteristics, conditions and preferences, what should I expect will happen to me?"
 - "What are my options and what are the potential benefits and harms of those options?"
 - "What can I do to improve the outcomes that are most important to me?"
 - "How can clinicians and the care delivery systems they work in help me make the best decisions about my health and healthcare?"

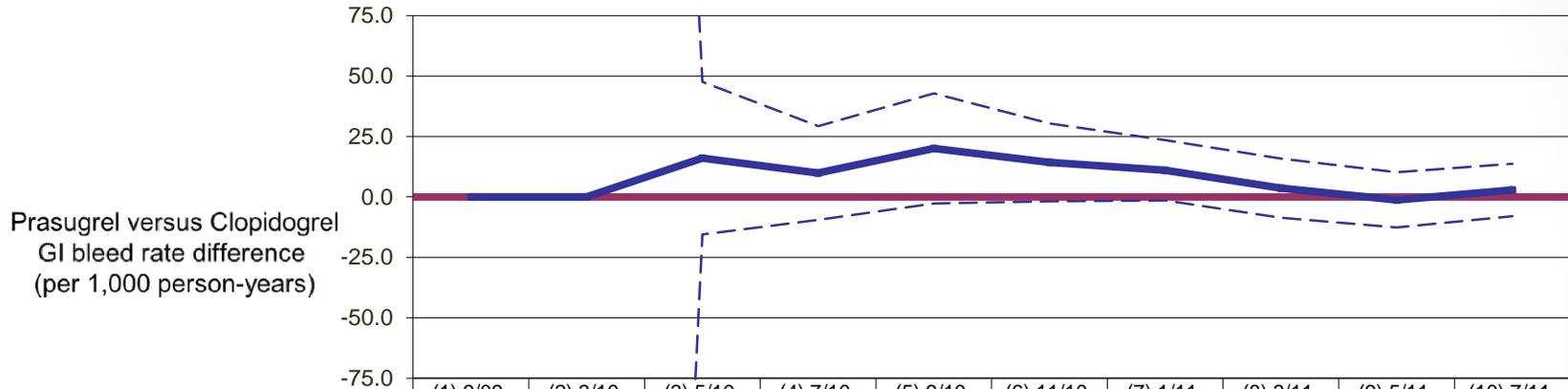
Making decisions about treatments

- At the policy level:
 - FDA market approval for drugs and devices
 - FDA's Sentinel Initiative for post-approval safety surveillance
 - Payer reimbursement (CMS, medical/pharmacy insurance)
- At the patient level (PCORI's mandate):
 - "Given my personal characteristics, conditions and preferences, what should I expect will happen to me?"
 - "What are my options and what are the potential benefits and harms of those options?"
 - "What can I do to improve the outcomes that are most important to me?"
 - "How can clinicians and the care delivery systems they work in help me make the best decisions about my health and healthcare?"

Example: Post-approval monitoring

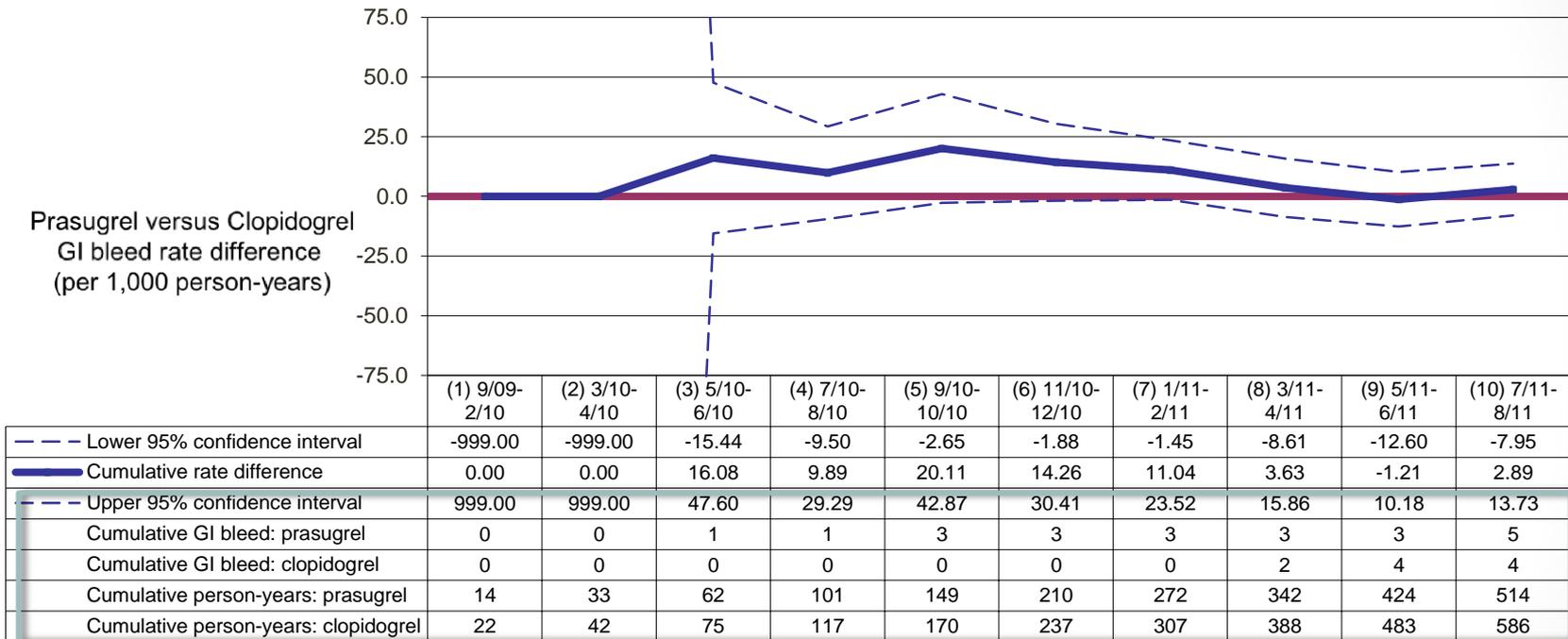
- We want to monitor the safety of a new anti-platelet, prasugrel, as compared to clopidogrel, an older drug with similar indications.
- TRITON-TIMI 38 showed prasugrel reduced cardiovascular death, and non-fatal cardiovascular events.
- Trial also showed a potentially higher risk of bleeding.
- We plan a sequential cohort observational study in healthcare claims data (NJ & PA Medicare) to monitor the bleeding risk.
- We assume treatment groups can be made comparable via matching or weighting in sequential cohorts.
- When is the evidence sufficient to raise an alert?
- How bad does the bleed risk need to be to overcome the improved cardiovascular effects?

Sequential cohort design



	(1) 9/09-2/10	(2) 3/10-4/10	(3) 5/10-6/10	(4) 7/10-8/10	(5) 9/10-10/10	(6) 11/10-12/10	(7) 1/11-2/11	(8) 3/11-4/11	(9) 5/11-6/11	(10) 7/11-8/11
--- Lower 95% confidence interval	-999.00	-999.00	-15.44	-9.50	-2.65	-1.88	-1.45	-8.61	-12.60	-7.95
— Cumulative rate difference	0.00	0.00	16.08	9.89	20.11	14.26	11.04	3.63	-1.21	2.89
--- Upper 95% confidence interval	999.00	999.00	47.60	29.29	42.87	30.41	23.52	15.86	10.18	13.73
Cumulative GI bleed: prasugrel	0	0	1	1	3	3	3	3	3	5
Cumulative GI bleed: clopidogrel	0	0	0	0	0	0	0	2	4	4
Cumulative person-years: prasugrel	14	33	62	101	149	210	272	342	424	514
Cumulative person-years: clopidogrel	22	42	75	117	170	237	307	388	483	586

Sequential cohort design

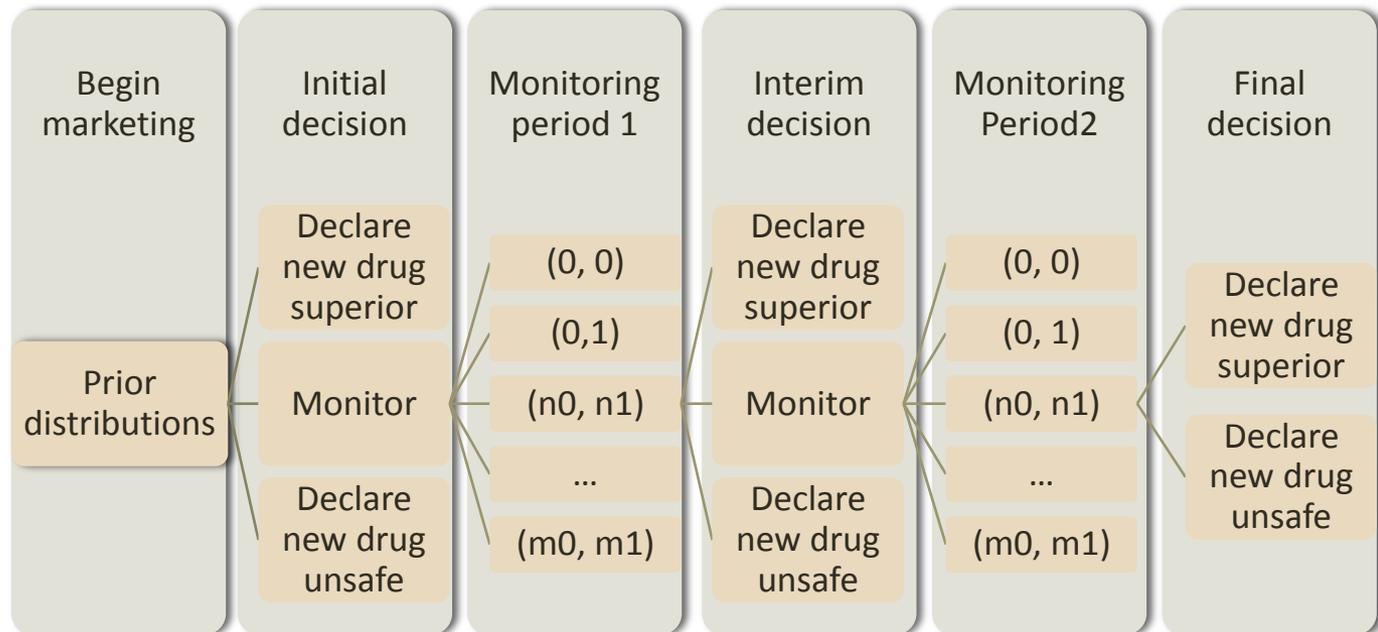


- Every 6 months, we get new data, and we identify initiators of either drug.
- Within the new initiators, we estimate a PS and perform PS matching.
- We then collect person-time and events within all matched patients (initiating in the current period and in previous periods).
- When do we decide that the safety-effectiveness profile for one drug exceeds the other?

Decision analysis

- Value of monitoring = improved treatment decisions and improved health outcomes for future patients.
- Cost of continued monitoring* = exposing current patients to unsafe drug.

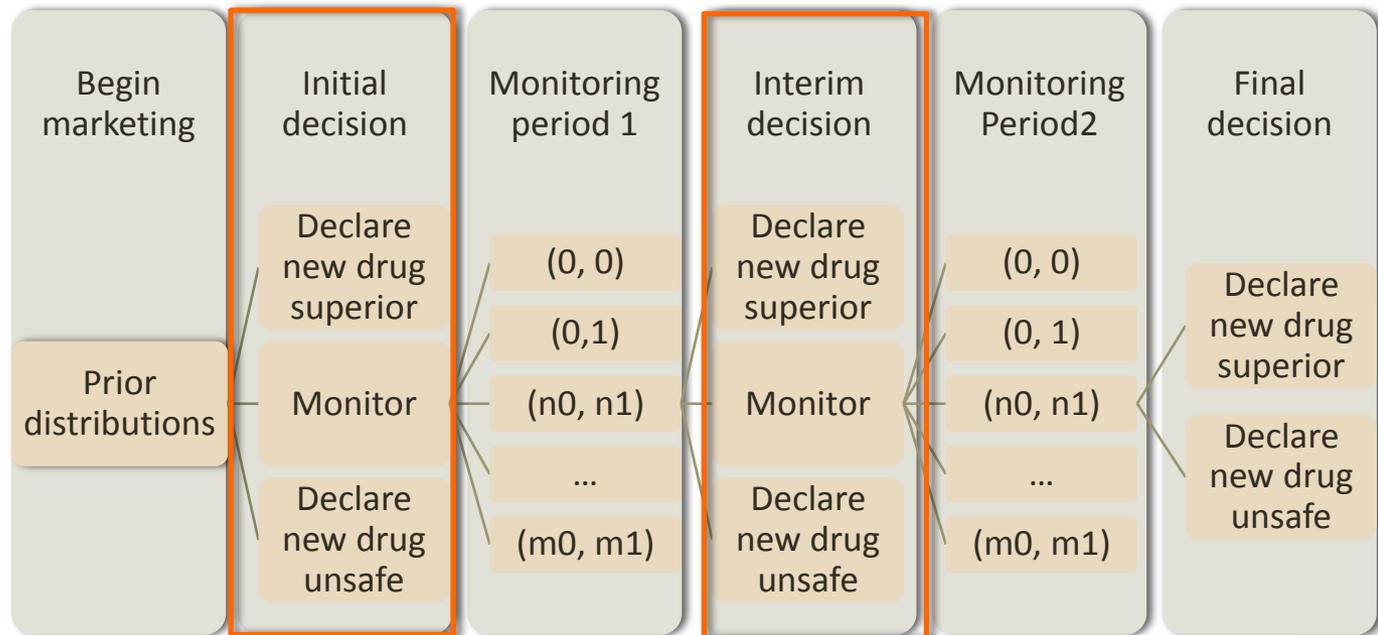
* Continued monitoring is equal to NOT making a decision



Decision analysis

- Value of monitoring = improved treatment decisions and improved health outcomes for future patients.
- Cost of continued monitoring* = exposing current patients to unsafe drug.

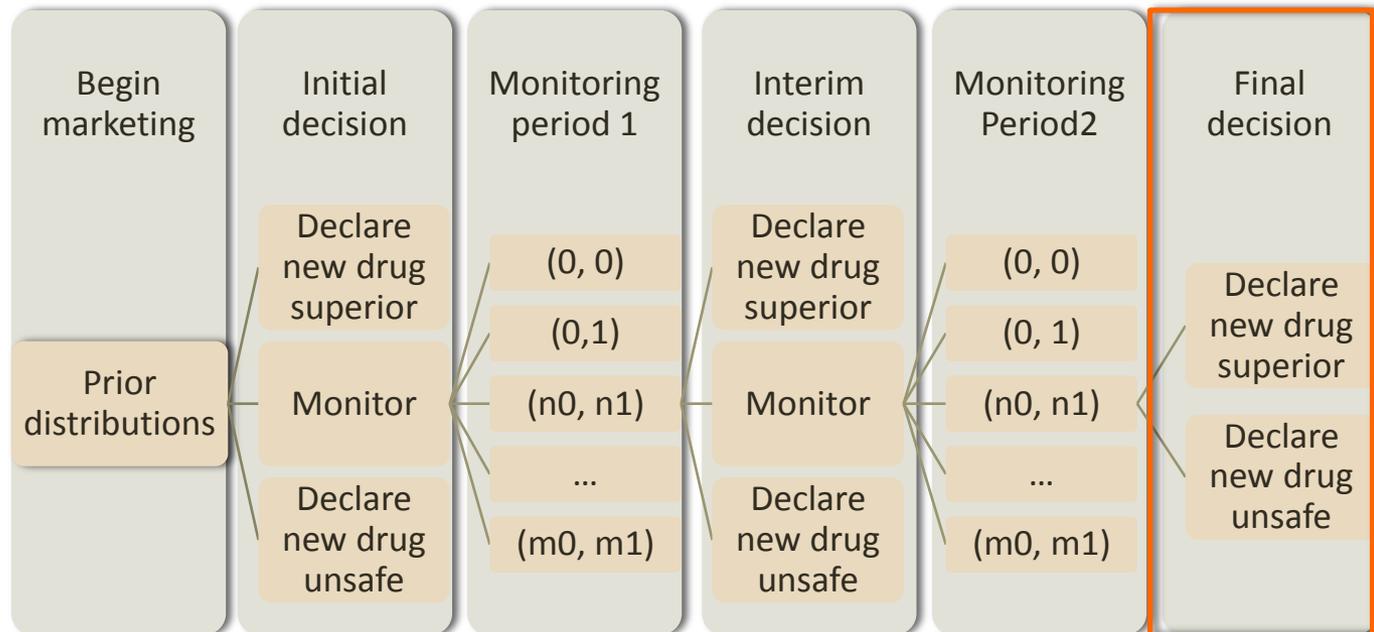
* Continued monitoring is equal to NOT making a decision



Decision analysis

- Value of monitoring = improved treatment decisions and improved health outcomes for future patients.
- Cost of continued monitoring* = exposing current patients to unsafe drug.

* Continued monitoring is equal to NOT making a decision



Past work

- Similar problems have been considered in the context of clinical trials.
 - Thompson (1933) considered adaptively assigning patients to treatment arms (Two-armed bandit).
 - Frequentist approaches to group sequential trials that account for multiple looks at the data:
 - Pocock (1977, 1982), O'Brien and Fleming (1979), Wang and Tsatis (1987), Muller and Schafer (2001), etc...
 - Bayesian approaches utilize decision theory for adaptive trial design:
 - Berry and Ho (1988), Lewis and Berry (1994), Stallard et al. (1999), Chen and Sheng (2005), etc...
- All prior approaches minimize loss functions within constraints of the pre-specified Type I error rate.

Our approach

- Our goal is to minimize the expected number of events over the entire population horizon.
- We calculate the expected events under each potential decision.
 - We assume that future decisions (after monitoring) will be made by again calculating expected events under each option, given observations from monitoring.
 - We assume that the decision to declare one drug superior will result in all patients being switched to that drug.
- Events on each treatment follow a Poisson distribution:

$$Y_a \sim \text{Poisson}(t\lambda_a)$$

- Gamma prior distributions on the rate parameters are updated after each monitoring period

$$\lambda_a \sim \text{Gamma}(\alpha_a, \beta_a)$$

Our approach

- Our goal is to minimize the expected number of events over the entire population horizon.
- We calculate the expected events under each potential decision.
 - We assume that future decisions (after monitoring) will be made by again calculating expected events under each option, given observations from monitoring.
 - We assume that the decision to declare one drug superior will result in all patients being switched to that drug.

- Events on each treatment follow a Poisson distribution:

$$Y_a \sim \text{Poisson}(t\lambda_a)$$

- Gamma prior distributions on the rate parameters are updated after each monitoring period

$$\lambda_a \sim \text{Gamma}(\alpha_a, \beta_a)$$

Our approach

- Our goal is to minimize the expected number of events over the entire population horizon.
- We calculate the expected events under each potential decision.
 - We assume that future decisions (after monitoring) will be made by again calculating expected events under each option, given observations from monitoring.
 - We assume that the decision to declare one drug superior will result in all patients being switched to that drug.
- Events on each treatment follow a Poisson distribution:

$$Y_a \sim \text{Poisson}(t\lambda_a)$$

- Gamma prior distributions on the rate parameters are updated after each monitoring period

$$\lambda_a \sim \text{Gamma}(\alpha_a, \beta_a)$$

Including information on effectiveness

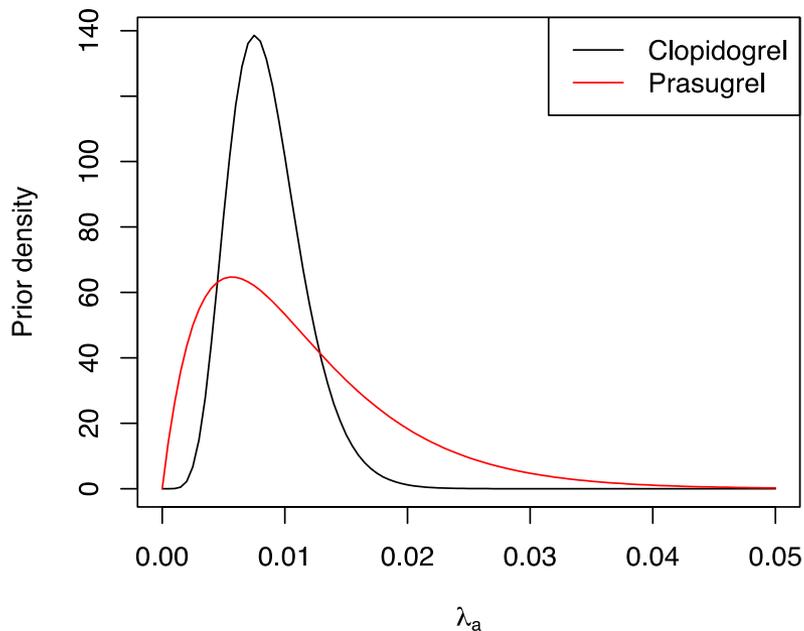
- We assume that the relative rates of effectiveness outcomes are known from RCTs
- We also must assume some relative preference weights for each outcome:
 - We base ours on the proportion of remaining life lost due to disability or death resulting from the event.
 - “Health losses”
 - Death gets a weight of 1.0
 - Could also use QALYs, DALYs, or any other weighting scheme.

Decision analytic input

Input	Value
Person-time prasugrel exposed during each monitoring period	13 person-years
Total person-time during treatment time horizon.	25,000 person-years
Maximum market-share for prasugrel	15%
Person-time to prasugrel reaching maximum market share	12,500 years
Prior for GI bleed rate, clopidogrel	Gamma(8, 930)
Prior for GI bleed rate, prasugrel	Gamma(2, 176)
Priors for additional safety and efficacy outcomes (rate per 100 person-years), clopidogrel	Non-fatal MI: 5.93 Non-fatal stroke: 1.14 Death: 1.27
Relative rates of safety outcomes, prasugrel versus clopidogrel	Non-fatal MI: 0.76 Non-fatal stroke: 1.02 Death: 0.95
Health loss weights (expressed as fractions of life expectancy)	GI bleed: 0.001 Non-fatal MI: 0.09 Non-fatal stroke: 0.25 Death: 1

Initial decision

- Prior to doing any monitoring, we calculate expected health losses under each potential decision.



Decision	Expected health losses
Declare prasugrel superior	475.32
Declare clopidogrel superior	521.71
Monitor	480.36

Sensitivity analyses

	prasugrel	monitor	clopidogrel	
Base case	475.32	480.36	521.71	
Person-time assumptions	1-way sensitivity analyses			
	Total person-time: increased to 50,000 PY	950.64	955.68	1043.43
	: reduced to 10,000 PY	190.13	195.17	208.69
	Prasugrel-exposed PY per cycle: increased to 100 PY	475.32	485.13	521.71
	Maximum prasugrel market-share increased to 50%	475.32	478.69	521.71
	Time (PY) to maximum prasugrel uptake: reduced to 2,500	475.32	477.03	521.71
	: increased to 25,000	475.32	477.82	521.71
Other outcome assumptions	Prasugrel market-share set to constant 15%	475.32	475.39	521.71
	Prasugrel bleed rate: increased to match upper 95% CL of HR	475.40	480.43	521.71
	: decreased to match lower 95% CL of HR	475.26	480.30	521.71
	Relative rate of non-fatal MI increased to 0.85	487.34	491.07	521.71
	Relative rate of non-fatal stroke increased to 1.45	505.93	507.64	521.71
	Relative rate of death increased to 1.16	541.85	521.14	521.71
	GI bleed weight increased to 0.007	477.14	482.13	523.09
Multi-way sensitivity analyses	Relative rates of MI, stroke and death set to 1	521.78	521.70	521.71
	Relative rates of stroke and death set to 1.	489.74	493.21	521.71
	Health loss weight for GI bleed at maximum, MI and stroke at minimum	363.26	366.44	392.54
Subgroup analysis				
Contra-indicated population	561.79	555.45	555.48	

Sensitivity analyses

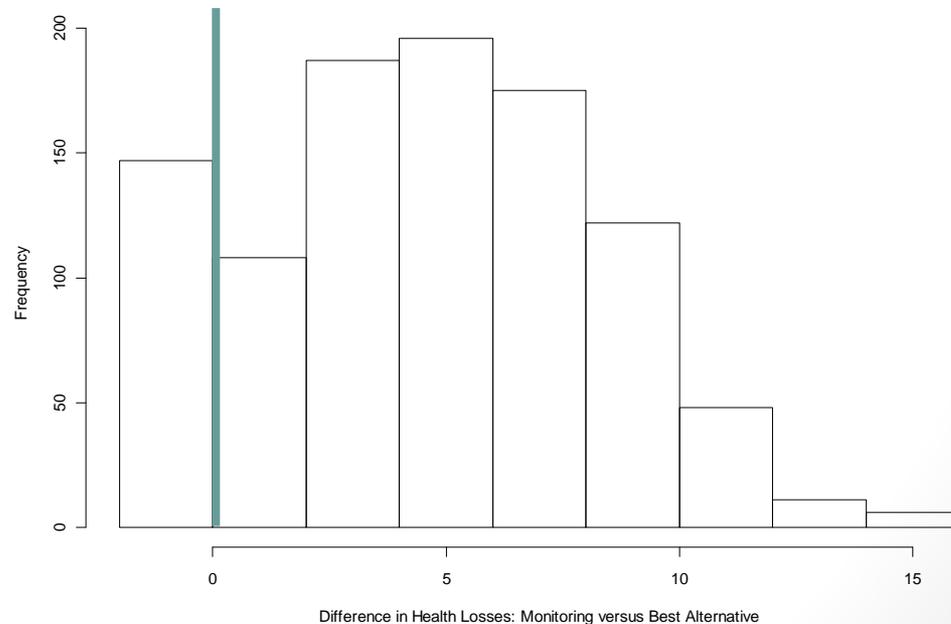
	prasugrel	monitor	clopidogrel	
Base case	475.32	480.36	521.71	
1-way sensitivity analyses				
Person-time assumptions	Total person-time: increased to 50,000 PY	950.64	955.68	1043.43
	: reduced to 10,000 PY	190.13	195.17	208.69
	Prasugrel-exposed PY per cycle: increased to 100 PY	475.32	485.13	521.71
	Maximum prasugrel market-share increased to 50%	475.32	478.69	521.71
	Time (PY) to maximum prasugrel uptake: reduced to 2,500	475.32	477.03	521.71
	: increased to 25,000	475.32	477.82	521.71
	Prasugrel market-share set to constant 15%	475.32	475.39	521.71
	Prasugrel bleed rate: increased to match upper 95% CL of HR	475.40	480.43	521.71
	: decreased to match lower 95% CL of HR	475.26	480.30	521.71
	Relative rate of non-fatal MI increased to 0.85	487.34	491.07	521.71
Other outcome assumptions	Relative rate of non-fatal stroke increased to 1.45	505.93	507.64	521.71
	Relative rate of death increased 0.95 -> 1.16	541.85	521.14	521.71
	GI bleed weight increased to 0.007	477.14	482.13	523.09
	Multi-way sensitivity analyses			
Relative rates of death				
Relative rates of health loss				
Health loss weights at minimum				
Subgroup analyses				
Contra-indications				

Death is extremely influential because of its high health loss weight. However, the effect on death is imprecise in the RCTs (RR= 0.95; 0.78 – 1.16)!

Health loss weights:
 GI bleed: 0.001
 Non-fatal MI: 0.09
 Non-fatal stroke: 0.25
Death: 1

Monte Carlo simulation

- In order to account for the uncertainty in relative rates of stroke, MI, and death, we sampled from their distributions based on randomized trial evidence.
- Using each set of sampled relative rates, we recalculated expected health losses under each strategy.



Sequential monitoring

- We identified patients from a commercial health insurer claims database.
 - Initiating prasugrel or clopidogrel following hospitalization for MI or ischemic heart disease with PCI.
 - 180 days enrollment before index exposure.
 - No exposure to either drug in 180 days before index.
 - Covariates were assessed during 180 days before index.
- During each monitoring period:
 - Variable ratio PS matching of exposure groups
 - Outcomes in the 180 days following exposure.
 - Updated priors and reevaluated the decision.

Sequential decision analyses

Period	Duration of period (months)	Clopidogrel initiators					Prasugrel initiators					Decision after monitoring
		Person-years	GI bleeds	strokes	MIs	Deaths (all-cause)	Person-years	GI bleeds	strokes	MIs	Deaths (all-cause)	
1	6	481	7	8	76	9	14	0	0	2	1	Treat PSG
2	2	321	0	7	54	3	20	0	1	4	0	Treat PSG
3	2	369	1	4	35	1	30	1	1	2	0	Treat PSG
4	2	412	1	2	31	5	41	0	0	5	0	Treat PSG
5	2	434	3	8	48	0	50	2	0	3	0	Treat PSG
6	2	465	1	4	32	12	64	0	0	4	0	Treat PSG
7	2	448	5	0	27	8	65	0	0	3	0	Treat PSG
8	2	482	18	25	44	9	73	0	0	6	0	Treat PSG
9	2	494	9	0	41	0	85	0	0	10	1	Treat PSG
10	2	498	10	2	28	15	93	2	1	8	1	Treat PSG

- The preference for prasugrel predicted by the initial decision analysis was maintained after each of monitoring period.

Benefits of the approach

- Requires upfront specification of preferences regarding values of benefits and risks.
- Minimizes overall health losses by incorporating multiple safety and efficacy outcomes.
- Assigns equal weight to outcomes of current and future patients with indications for treatment.

Limitations of the approach

- Our approach makes several simplifying assumptions:
 - Unbiased estimates of comparative safety can be obtained through the use of adjustment.
 - Efficacy outcomes will not be updated during monitoring, and thus underestimates the value of monitoring.
 - Patient preferences and treatment effects are homogeneous.

Conclusions

- VOI analyses highlight the decision-making process in safety monitoring activities.
- Input on other questions should also be considered:
 - Uncertainty in prior assumptions
 - Heterogeneity in patient population
 - Reduced market competition and treatment options if safety alert is issued

Thanks

- My collaborators:
 - Amanda Patrick
 - Robert Glynn
 - Milton Weinstein
 - Sebastian Schneeweiss
- Questions?
- Suggestions???