Healthcare Database Healthcare Data Science to Quantify Adverse Health Effects

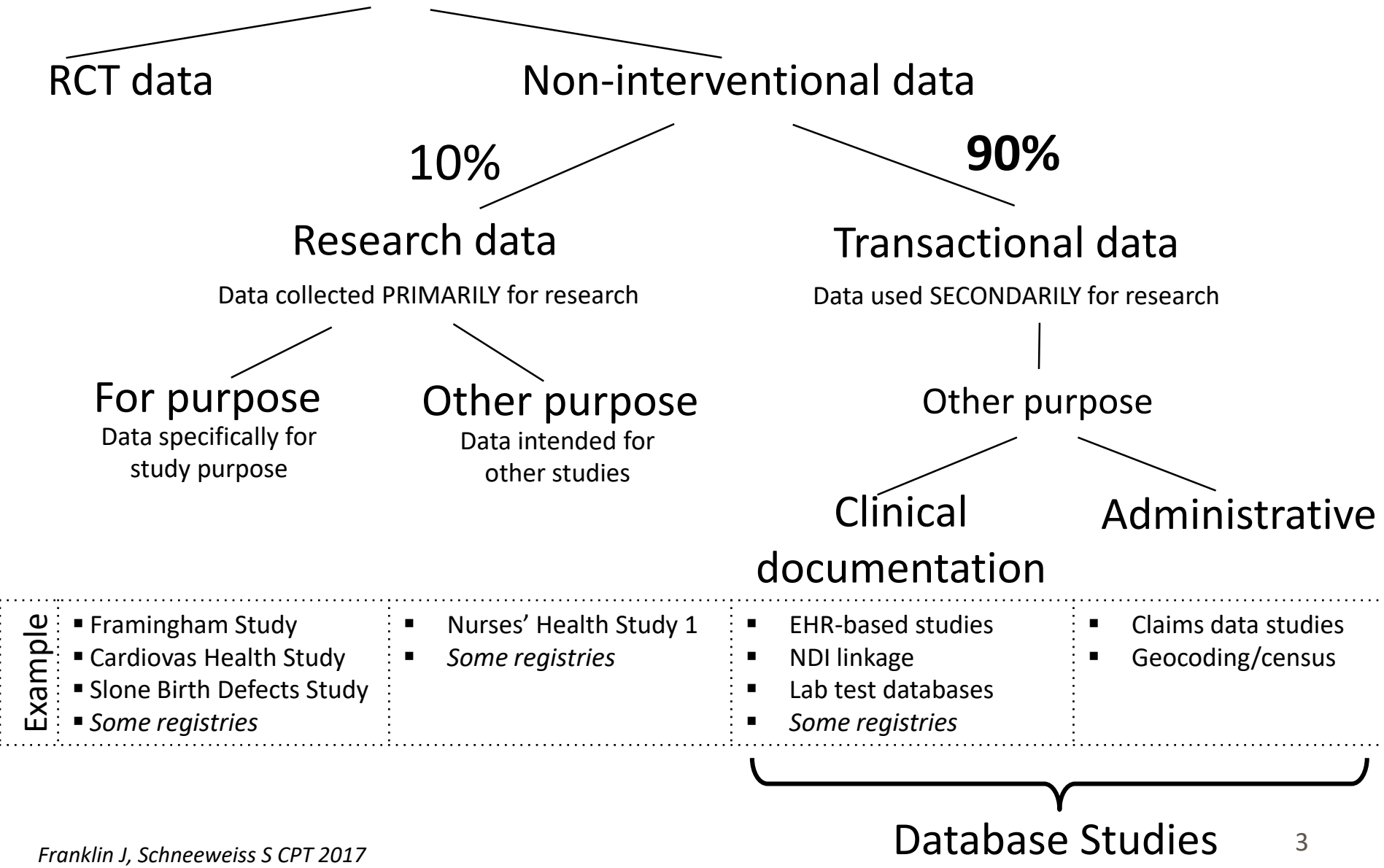
Sebastian Schneeweiss, MD, ScD

Division of Pharmacoepidemiology and Pharmacoeconomics,
Dept. of Medicine, Brigham & Women's Hospital/ Harvard Medical School

Conflicts of Interest

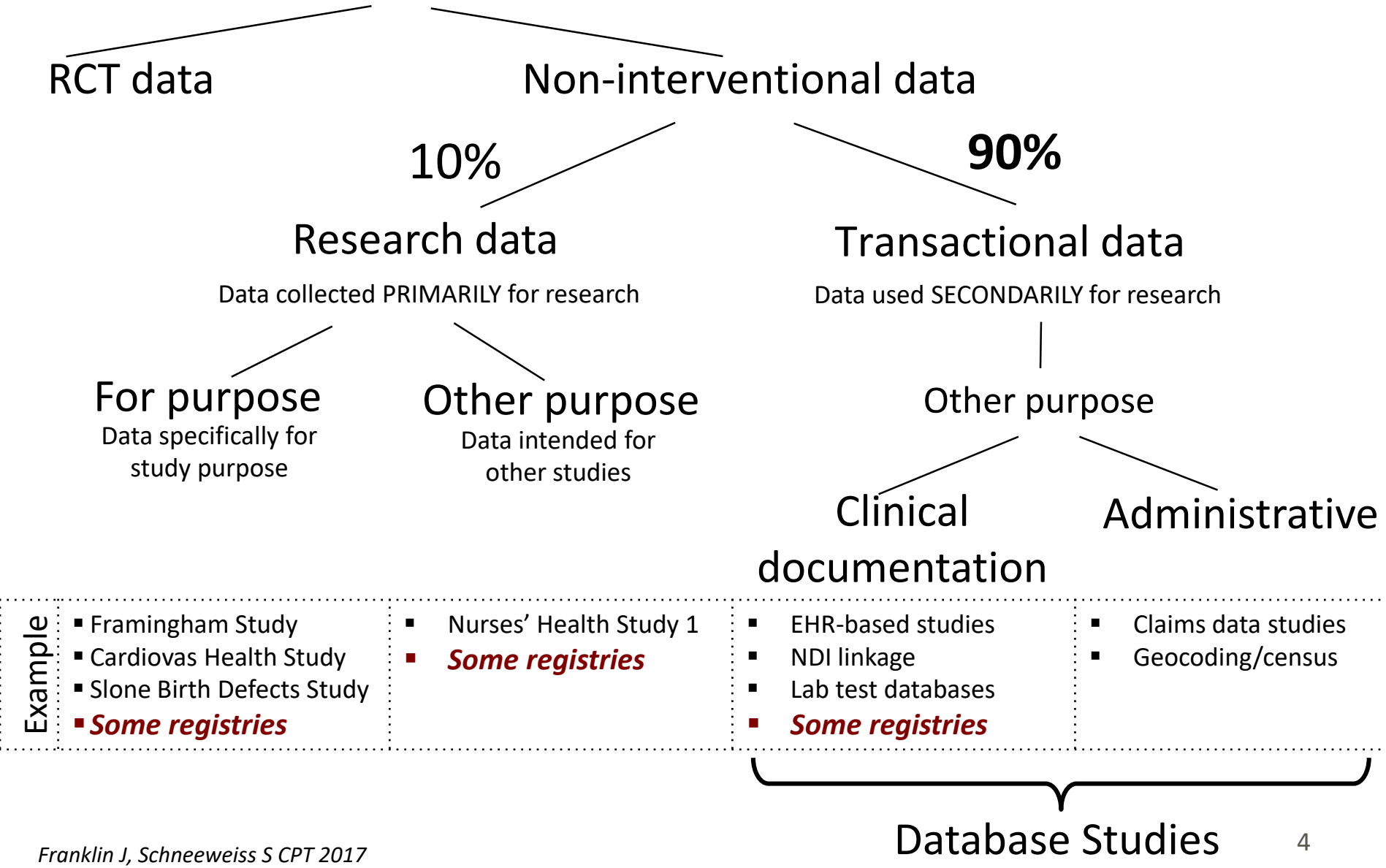
PI, Harvard-Brigham & Women's Hospital Drug Safety Research Center (FDA); Co-Chair, Methods Core of the FDA Sentinel System; Consulting in past year: WHISCON LLC, Aetion Inc. (incl. equity); PI of research contracts to the Brigham & Women's Hospital: Bayer, Genentech, Boehringer Ingelheim; Grants/contracts from NIH, AHRQ, PCORI, FDA, IMI, Arnold Foundation; Advising FDA, EMA, PCORI, PMDA, Health Canada.

Effectiveness Research with Healthcare Databases



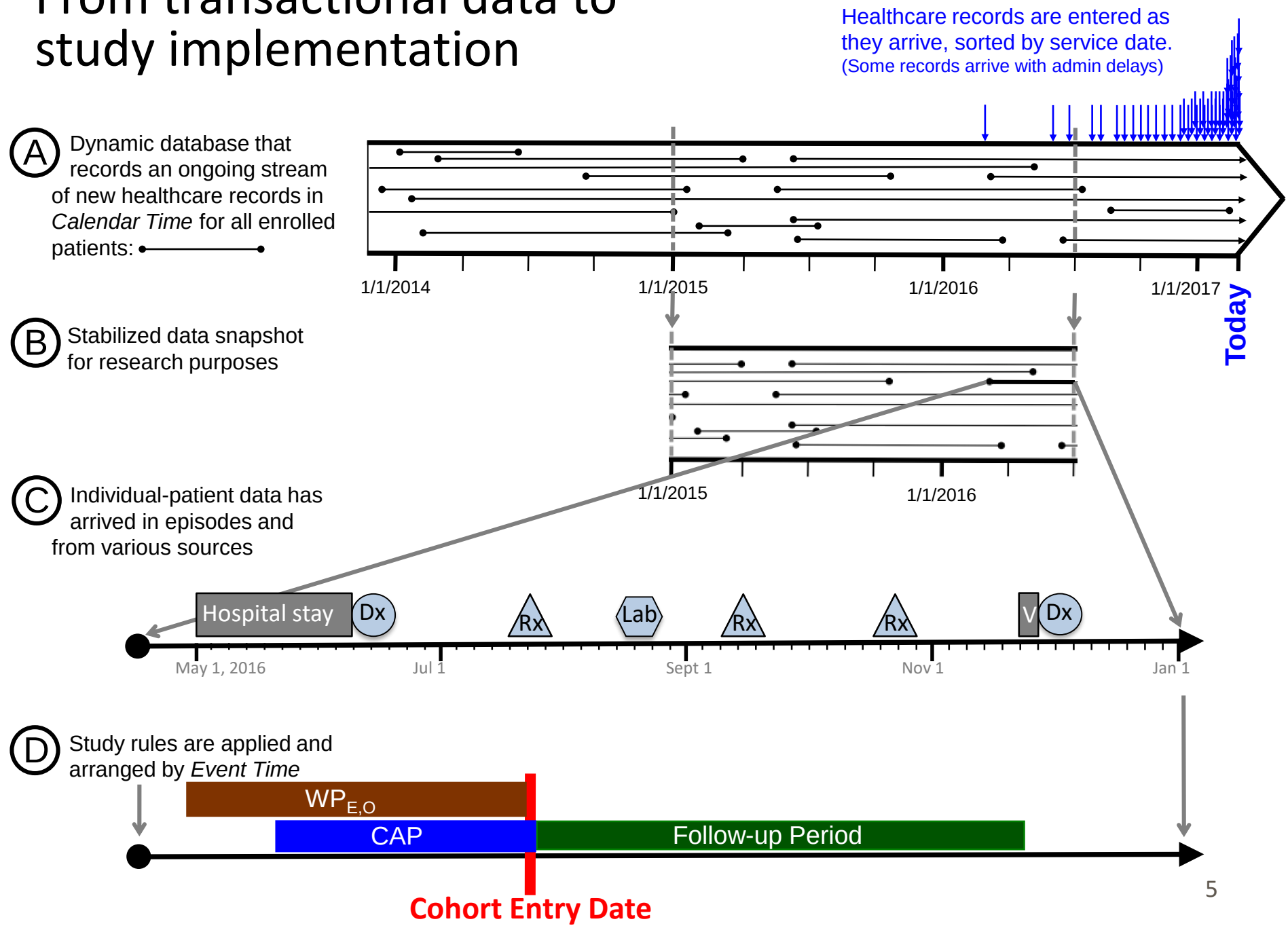
Franklin J, Schneeweiss S CPT 2017

Effectiveness Research with Healthcare Databases



Franklin J, Schneeweiss S CPT 2017

From transactional data to study implementation



In-hospital safety examples **blinded** with respect to RCT findings:

followed by

Database Study → **RCT**

The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 FEBRUARY 21, 2008 VOL. 358 NO. 8

Aprotinin during Coronary-Artery Bypass Grafting and Risk of Death

Sebastian Schneeweiss, M.D., Sc.D., John D. Seeger, Pharm.D., Dr.P.H., Joan Landon, M.P.H., and Alexander M. Walker, M.D., Dr.P.H.

Risk of death (7d)

HR = 1.78 (1.56 -2.02)

Outcome	Any Amount of Aprotinin (N=33,517)	Any Amount of Aminocaproic Acid (N=44,682)	Any Amount of Study Drug		Low or High Amount of Study Drug
			Unadjusted	Adjusted	Adjusted
	no. of patients (%)		*relative risk (95% CI)*		
In-hospital death from any cause	1512 (4.5)	1101 (2.5)	1.83 (1.70–1.98)	1.64 (1.50–1.78)	1.50 (1.36–1.66)
In-hospital death from any cause within 7 days after CABG	631 (1.9)	435 (1.0)	1.93 (1.71–2.18)	1.78 (1.56–2.02)	1.64 (1.41–1.91)

The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 MAY 29, 2008 VOL. 358 NO. 22

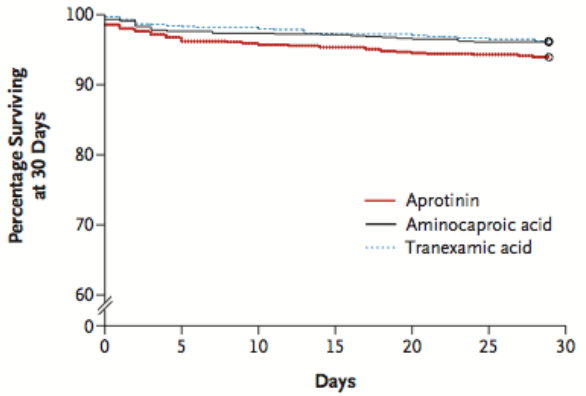
A Comparison of Aprotinin and Lysine Analogues in High-Risk Cardiac Surgery

Dean A. Fergusson, M.H.A., [illegible], M.D., Stephen Fremes, M.D., Charles MacAdams, [illegible], Peter C. Duke, M.D., Ramiro Arellano, M.D., M.Sc., [illegible] Côté, M.D., Jacek Karski, M.D., Raymond Martineau, M.D., [illegible], M.Sc., George Wells, Ph.D., Jennifer Clin[illegible]nvestigators†

BART

Risk of death (30 d)

HR = 1.53 (1.06 -2.22)



No. at Risk							
Aprotinin	779	753	747	742	737	734	732
Aminocaproic acid	780	761	759	757	753	749	749
Tranexamic acid	769	757	755	748	747	743	749

CV safety example **blinded** with respect to RCT findings:

Database Study *followed by* **RCT**

ARTHRITIS & RHEUMATOLOGY

Cardiovascular Safety of Tocilizumab Versus Tumor Necrosis Factor Inhibitors in Patients With Rheumatoid Arthritis

A Multi-Database Cohort Study

Seoyoung C. Kim,[1] Daniel H. Solomon,[1] James R. Rogers,[1] Sara Gale,[2] Micki Klearman,[2] Khaled Sarsour,[2] and Sebastian Schneeweiss[1]

Risk of composite CV outcome

HR = 0.85 (0.61-1.19)

	TCZ				
	No. of subjects	No. of events	Person-years	IR (95% CI)†	HR (95% CI)
As-treated analysis					
Composite cardiovascular events					
Medicare	2,531	17	1,841	0.92 (0.56–1.44)	0.70 (0.40–1.24)
PharMetrics	2,614	10	2,061	0.49 (0.25–0.86)	1.00 (0.45–2.22)
MarketScan	4,073	9	2,999	0.30 (0.15–0.55)	1.03 (0.46–2.34)
Combined	9,218	36	6,901	0.52 (0.37–0.71)	0.84 (0.56–1.26)‡

ABSTRACT NUMBER: 3L

Comparative Cardiovascular Safety of Tocilizumab Vs Etanercept in Rheumatoid Arthritis: Results of a Randomized, Parallel-Group, Multicenter, Noninferiority, Phase 4 Clinical Trial

ENTRACTE

Jon T. Giles[1], Naveed Sattar[2], Sherine E. Gabriel[3], Paul M. Ridker[4], Steffen Gay[5], Charl David Musselman[7], Laura Brockwell[6], Emma Shittu[6], Micki Klearman[7] and Thomas Fl

Risk of composite CV outcome

HR = 1.05 (0.77-1.43)

Etanercept N = 1542	Tocilizumab N = 1538	Tocilizumab vs Etanercept	
First Events, n	First Events, n	HR[a]	95% CI
78	83	1.05	0.77, 1.43

Effectiveness Example **blinded** with respect to RCT findings:

Database Study *followed by* **RCT**

American Diabetes Association. 77TH SCIENTIFIC SESSIONS | EXPERIENCE NEW HORIZONS IN DIABETES | SAN DIEGO, CA JUNE 9-13, 2017

Cardiovascular outcomes associated with canagliflozin versus other non-gliflozin antidiabetic drugs: population based cohort study

Elisabetta Patorno,[1] Allison B Goldfine,[2] Sebastian Schneeweiss,[1] Brer
Robert J Glynn,[1] Jun Liu,[1] Seoyoung C Kim[1,4]

BMJ

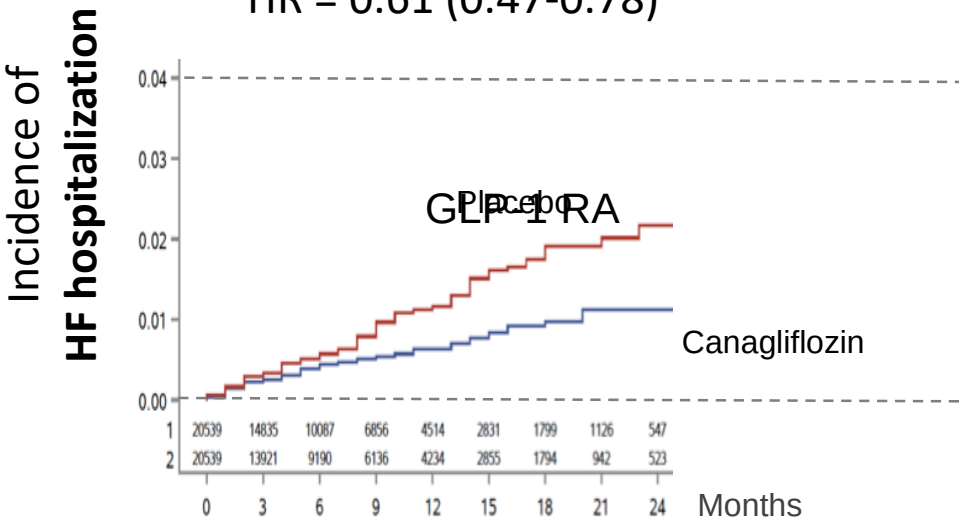
ORIGINAL ARTICLE

Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes

Bruce Neal, M.B., Ch.B., Ph.D., Vlado Perkovic, M.B., B.S., Ph.D., Kenneth W. Mahaffey, M.D., Dick de Zeeuw, M.D., Ph.D., Greg Fulcher, M.D., Ngozi Erondu, M.D., Ph.D., Wayne Shaw, D.S.L., Gordon Law, Ph.D., Mehul Desai, M.D., and David R. Matthews, D.Phil., B.M., B.Ch., for the CANVAS Program Collaborative Group*

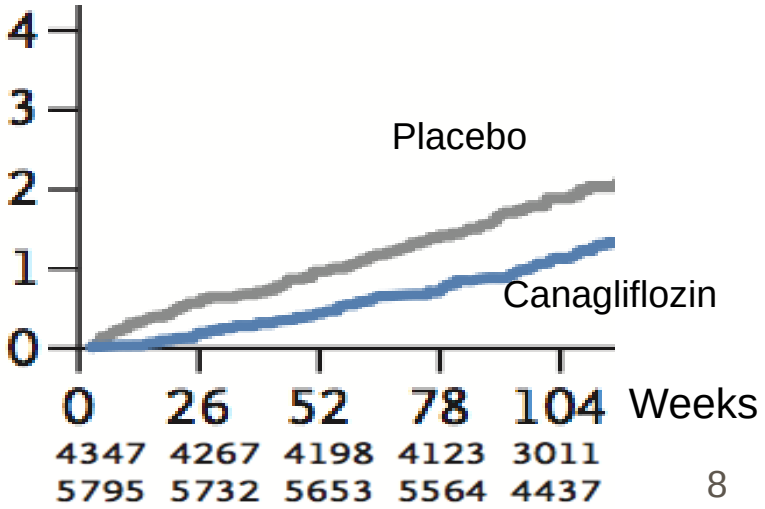
Prevention of heart failure hospitalization

HR = 0.61 (0.47-0.78)



Prevention of heart failure hospitalization

HR = 0.67 (0.52-0.87)



Safety Example **after** RCT findings were released: **Confirming signal**

RCT *followed by* **Database Study**

ORIGINAL ARTICLE

Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes

Bernard Zinman, M.D., Christoph Wanner, M.D., John M. Lachin, Sc.D., David Fitchett, M.D., Erich Bluhmki, Ph.D., Stefan Hantel, Ph.D., Michaela M..., ...s, Dr.P.H., Odd Erik Johansen, ...li C. Broedl, M.D., and Silvio E. Inzucchi, M.D., for the EMPA-REG OUTCOME Investigators

EMPA-REG

CORRESPONDENCE

Risk of Diabetic Ketoacidosis after Initiation of an SGLT2 Inhibitor

Michael Fralick, M.D.
Sebastian Schneeweiss, M.D., Sc.D.
Elisabetta Patorno, M.D., Dr.P.H.

Empagliflozin and risk of DKA

1 / 2,333 vs. 3 / 2,345

HR = 2.9 (0.4-20.0)

Table 2. Adverse Events.*

Event	Placebo (N=2333)	Empagliflozin, 10 mg (N=2345)	Empagliflozin, 25 mg (N=2342)	Pooled Empagliflozin (N=4687)
	number of patients (percent)			
Diabetic ketoacidosis¶¶	1 (<0.1)	3 (0.1)	1 (<0.1)	4 (0.1)

SGLT-2 and risk of DKA

26 / 38,045 vs. 55 / 38,045

HR = 2.2 (1.4-3.6)

Table 2. Primary and Other Outcomes.*

Days of Follow-up	DPP4 Inhibitor (N=38,045)		SGLT2 Inhibitor (N=38,045)	
	Diabetic Ketoacidosis	Hazard Ratio	Diabetic Ketoacidosis	Hazard Ratio (95% CI)
	no. of patients (rate per 1000 person-yr)		no. of patients (rate per 1000 person-yr)	
180 Days of follow-up†	26 (2.2)	1.0	55 (4.9)	2.2 (1.4–3.6)
60 Days of follow-up	13 (2.3)	1.0	31 (5.6)	2.5 (1.3–4.7)
30 Days of follow-up	10 (3.3)	1.0	22 (7.5)	2.3 (1.1–4.8)
180 Days of follow-up among patients not receiving insulin‡	9 (1.0)	1.0	21 (2.5)	2.5 (1.1–5.5)

Effectiveness Example **after** RCT findings were released:

RCT *followed by* → **Database Study**

The NEW ENGLAND JOURNAL of MEDICINE
ESTABLISHED IN 1812 SEPTEMBER 17, 2009 VOL. 361 NO. 12

Dabigatran versus Warfarin in Patients with Atrial Fibrillation

Stuart J. Connolly, M.D[illegible]R.C.P.C., D.Phil., John Eikelboom, M.D., Jonas[illegible]Sc., Paul A. Reilly, Ph.D., Ellison Themeles, B.A., Jeann[illegible]h.D., Denis Xavier, M.D., Jun Zhu, M.D., Rafael Diaz, M[illegible]oph Diener, M.D., Ph.D., Campbell D. Joyner, M.D., [illegible]tee and Investigators*

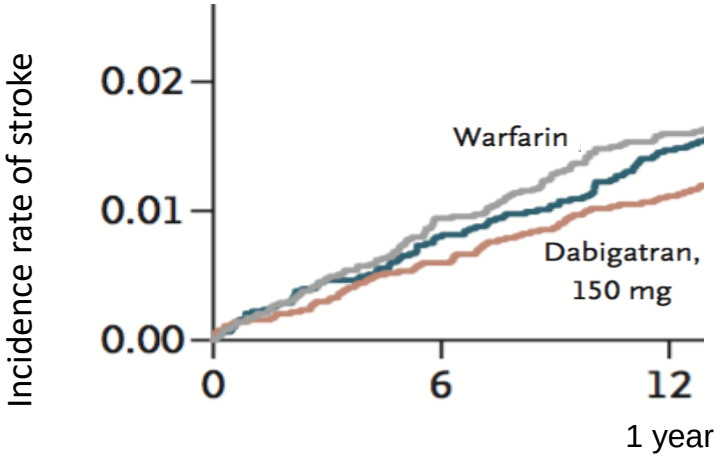
RE-LY

Thrombosis and Haemostasis
International Journal for Vascular Biology and Medicine

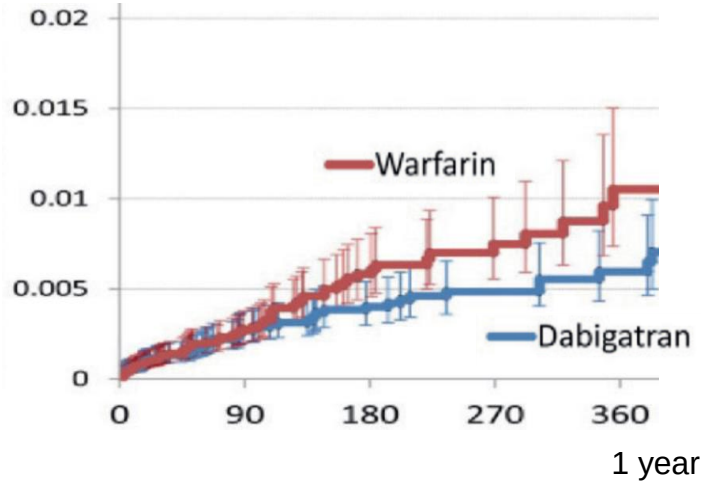
Safety and effectiveness of dabigatran and warfarin in routine care of patients with atrial fibrillation

John D. Seeger[1]; Katsiaryna Bykov[1]; Dorothee B. Bartels[2,3]; Krista Huybrechts[1]; Kristina Zint[2]; Sebastian Schneeweiss[1]

Stroke prevention
HR = 0.66 (0.53-0.82)



Stroke prevention
HR = 0.77 (0.54-1.09)



Key information components

- **Accurate assessment of Exposure:**
  - Completeness of repeated uses
  - Prescribing vs. dispensing vs. use of drugs

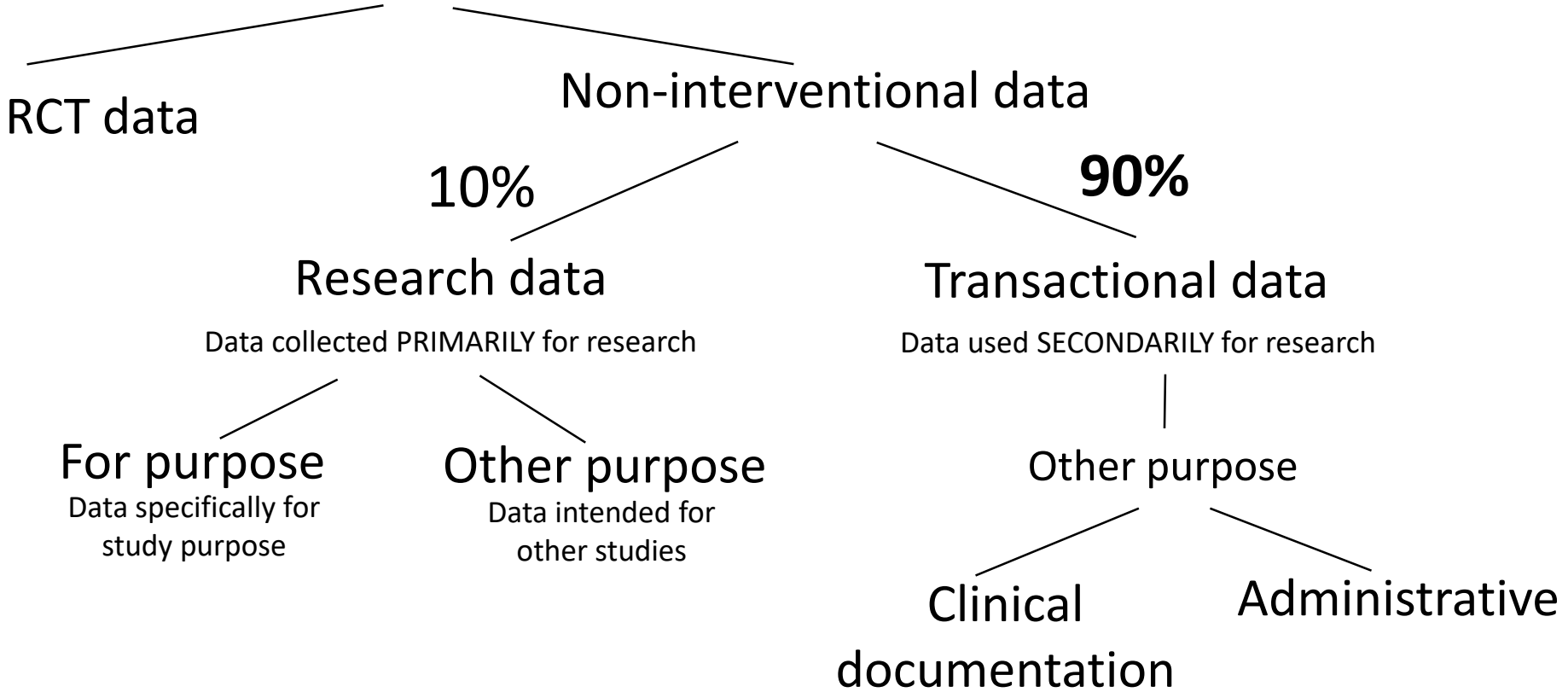
Interview
Pill counter

- **Accurate assessment of Outcome:**
  - High specificity of outcome assessment when estimating relative effect measures: risk ratio, rate ratio, hazard ratio
  - Reasonable sensitivity to preserve event counts
- **Complete assessment of Confounders:**
  - Reduced unobserved confounding
  - Pre-exposure measurement to avoid adjustment for intermediates

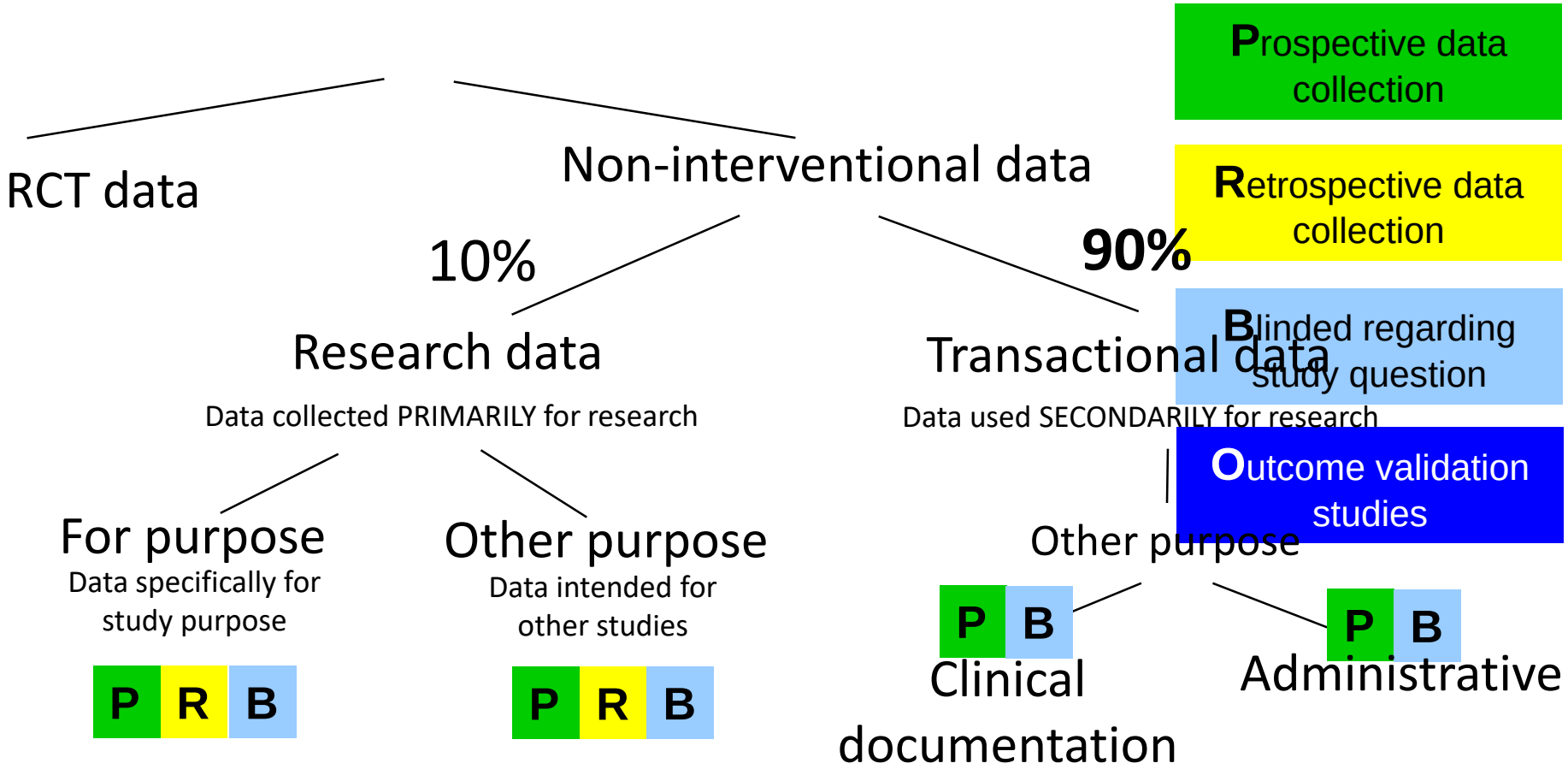
How were data generated?

What does that tell us about the quality of data?

For our study? (Fit-for-Purpose)



Example	▪ Framingham Study ▪ Cardiovas Health Study ▪ Slone Birth Defects Study ▪ *Some registries*	▪ Nurses' Health Study 1 ▪ *Some registries*	▪ EHR-based studies ▪ NDI linkage ▪ Lab test databases ▪ *Some registries*	▪ Claims data studies ▪ Geocoding/census
E				
O				
C				



Example	▪ Framingham Study ▪ Cardiovas Health Study ▪ Slone Birth Defects Study ▪ *Some registries*	▪ Nurses' Health Study 1 ▪ *Some registries*	▪ EHR-based studies ▪ NDI linkage ▪ Lab test databases ▪ *Some registries*	▪ Claims data studies ▪ Geocoding/census
E				
O	O	O	O	O
C				

Framingham Study (cohort)

Major: **Biennial examination** procedures with extensive examination + interview

Additional: NDI linkage

Drug exposure assessment	Current or past use of estrogen @ biennial exam; **No start date, no stop date**
Outcome assessment	Physician review of **clinical notes, hospital and physician records** and **death certificates**. New Q waves in ECG since last visit. Stroke confirmed by review panel w/ neurologists
Confounder assessment	Very detailed, pre-exposure
Population size	5k – 20k

Wilson PW, Garrison RJ, Castelli WP. Postmenopausal estrogen use, cigarette smoking, and cardiovascular morbidity in women over 50. The Framingham Study. *N Engl J Med.* Oct 24 1985;313(17):1038-1043

Nurses' Health Study (cohort)

Major: **Biennial self-administered questionnaires**

Additional: Endpoint validation with medical records; NDI linkage

Drug exposure assessment	"Are you currently taking any of the following medications at least once a week" **No start date, no stop date** (Consequences: Hernan et al)
Outcome assessment	Non-fatal events: permission for **medical records review** (exposure blinded) Fatal events: Family + Med Records + NDI linkage
Confounder assessment	Very detailed, pre-exposure
Population size	100k

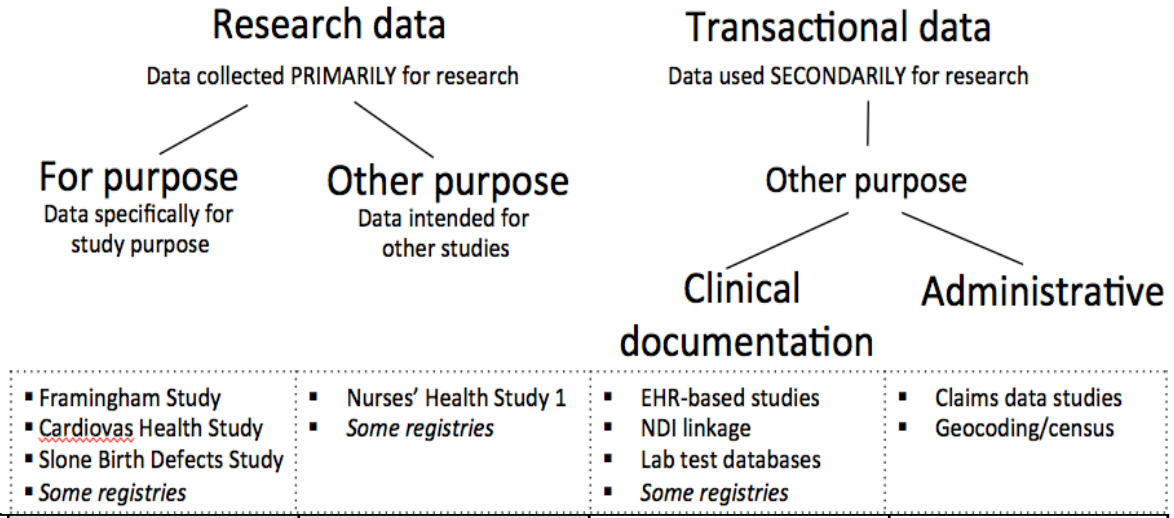
Michels KB, Rosner BA, Manson JE, et al. Prospective study of calcium channel blocker use, cardiovascular disease, and total mortality among hypertensive women: the Nurses' Health Study. *Circulation.* Apr 28 1998;97(16):1540-1548.
Stampfer MJ, Colditz GA, Willett WC, et al. Postmenopausal estrogen therapy and cardiovascular disease. Ten-year follow-up from the nurses' health study. *N Engl J Med.* Sep 12 1991;325(11):756-762.
Hernan MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology.* Nov 2008;19(6):766-779

Fundamental difference between primary vs. secondary data

Control over:	Primary (research) data: **Investigator defines measurements**	Secondary (transactional): **Business purpose defines measurement**
Which items will be measured	Targeted measurements for research study -> little unobserved factors	Information necessary to get the business done
How items will be measured	Measurement methods designed by investigator -> sufficient accuracy	Measurement good enough for business purpose
What **surveillance** will be in place to measure items?	Measurements actively scheduled -> high completeness	Measurements tied to healthcare encounters -> informative missingness (sicker patients with more encounters have more opportunity to have Dx recorded)

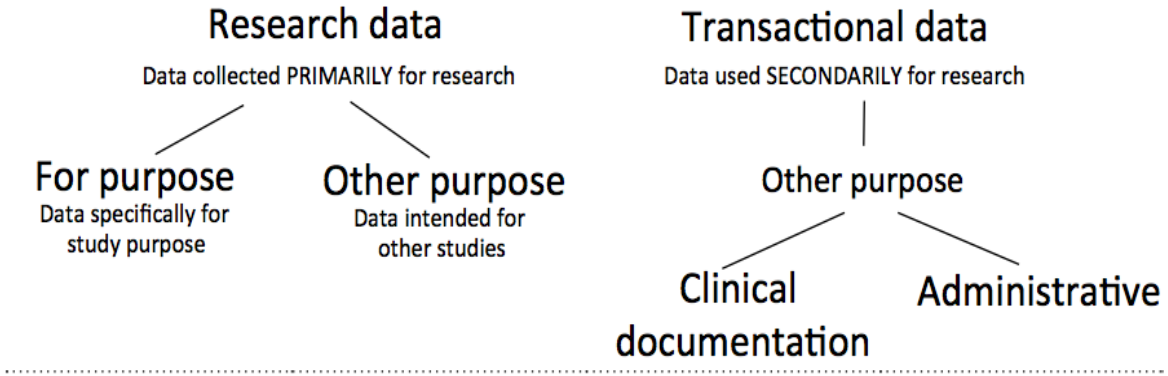
Secondary data work best if business interests are serendipitously aligned with research interests

Examples: Outcome assessment



	▪ Framingham Study ▪ Cardiovas Health Study ▪ Slone Birth Defects Study ▪ *Some registries*	▪ Nurses' Health Study 1 ▪ *Some registries*	▪ EHR-based studies ▪ NDI linkage ▪ Lab test databases ▪ *Some registries*	▪ Claims data studies ▪ Geocoding/census
Event surveillance				
Medial records review				
Death certificate				
Ultrasound				

Summary (Example)



Drug exposure assessment	C (A-)	C-	C-	A-
Confounder assessment	A	B+	A	B-
Outcome assessment	A	A	A-	B
Population size	**5k – 20k**	**100k**	**10m**	**100m**
0.1% exposed	**5-20**	**100**	**10k**	**100k**
1% exposed	**50-200**	**1k**	**100k**	**1m**

Conclusion

- There is no single perfect data source or study character
- Fit-for-purpose considerations
  - Exposure assessment
  - Endpoint assessment
  - Risk factors assessment before MRI exposure
- Clinical data do well
  - In detailed risk factor assessment
  - Outcome assessment
- Clinical data struggle:
  - Size
  - Prescription drug assessment