

What can be learned from observational patient-level databases?

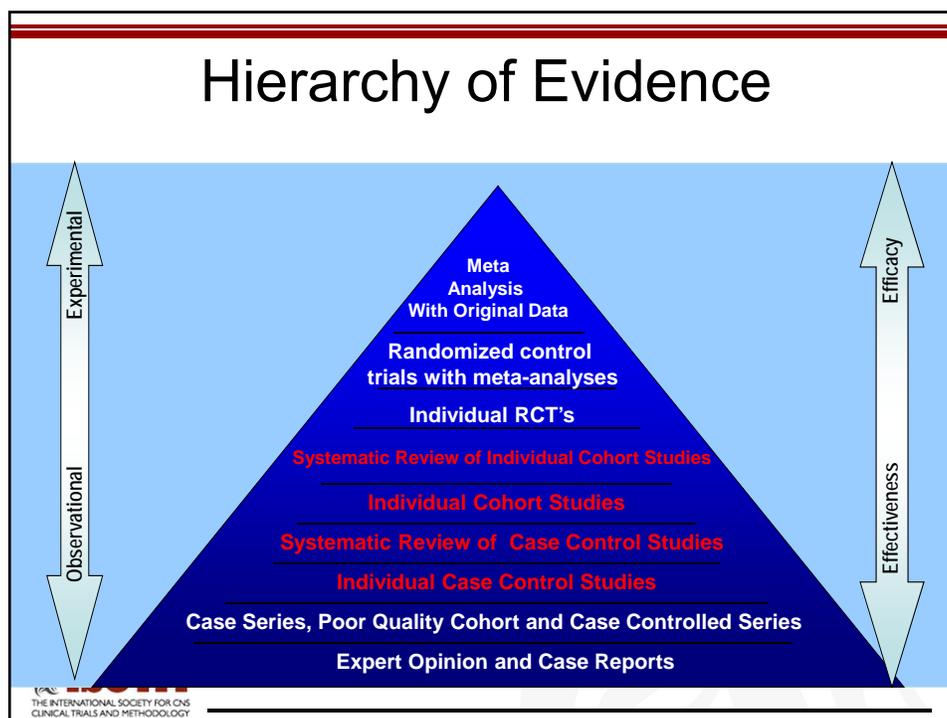
Paul Stang, PhD
Global Epidemiology
Janssen R&D



Disclosure

- Full-time employee of Janssen Research & Development
- Shareholder of Johnson & Johnson Stock





RCT vs. Observational Studies

- Both introduce biases
- Outcome: population-level mean effect
 - Observational studies: more efficiently move to individual effect estimates
- Database studies only useful if product on market and captured in that database
- *Observational studies: reflect the collective experience of clinicians in practice*
 - *Sensitive to data, method, & definitions*

NSAIDS and Prevention of Dementia Observational Data + / RCT data -

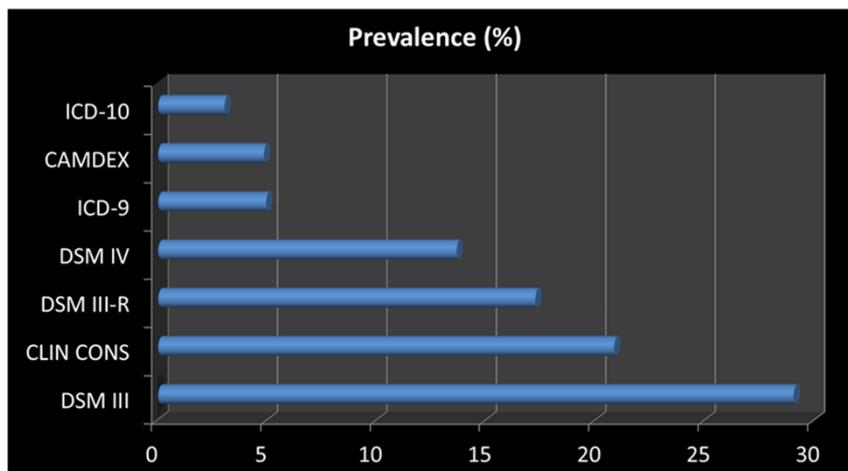
Study reference	Overall cohort	Duration NSAID use	AD cases	Risk ratio	95% confidence interval
PROSPECTIVE STUDIES					
Stewart et al. (1997)	1,686	≥2 years	81	0.40	0.19-0.84
in't Veld et al. (2001)	6,989	<2 years	4	0.65	0.33-1.29
		1-23 months	210	0.83	0.05-0.83
		<1 month	88	0.95	0.62-1.11
Breteler et al. (2002)	7,983	≥18 months	293	0.60	0.30-1.20
Zandi et al. (2002)	3,224	≥2 years	104	0.45	0.17-0.79
Cornelius et al. (2004)	1,301	NA	164	0.61	0.32-1.15
Haag et al. (2006)	6,992	≥2 years	582	0.65	0.40-1.06
Szekely et al. (2008a)	3,229	NA	321	0.63	0.45-0.88
Arvanitakis et al. (2008)	1,019	NA	209	1.19	0.87-1.62
Breitner et al. (2009)	2,736	NA	356	1.57	1.10-2.23
RETROSPECTIVE STUDIES					
Landi et al. (2003)	2,708	NA	269	0.43	0.23-0.82
Yip et al. (2005)	1,034	>6 months	61	0.64	0.24-0.98
Vlad et al. (2008)	246,199	>5 years	49,349	0.76	0.68-0.85
		>4 to ≤5 years		0.76	0.69-0.84
		>3 to ≤4 years		0.90	0.84-0.97
		>2 to ≤3 years		0.93	0.88-0.99
		>1 to ≤2 years		0.90	0.86-0.94
		≤1 year		0.98	0.95-1.00

NA: not applicable.

Front. Aging Neurosci., 21 May 2010 | <http://dx.doi.org/10.3389/fnagi.2010.00019>



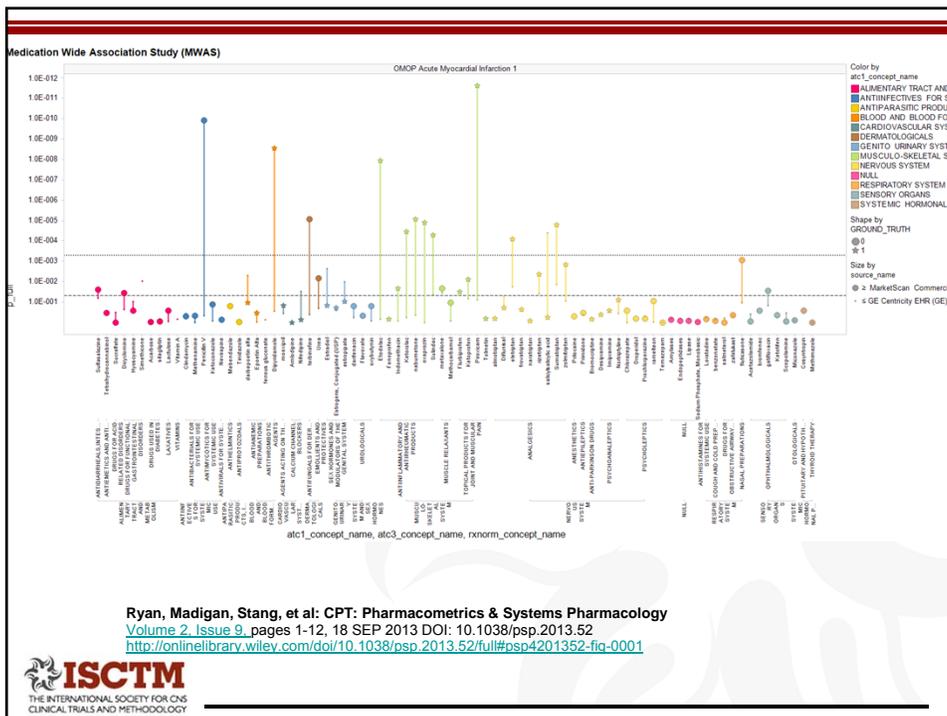
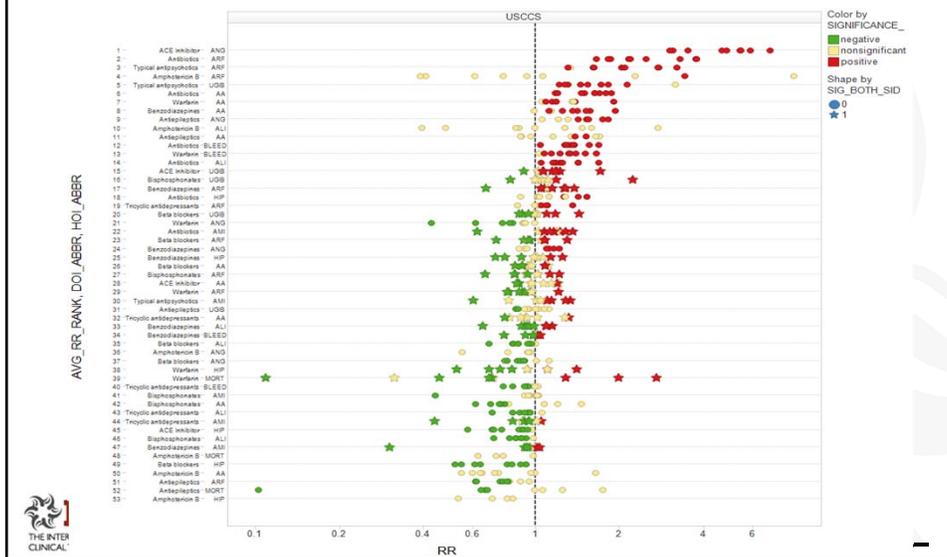
Which definition for Dementia?



Created from Erkinjuntti T, Masbye T, Steenhuis R, Hachinski V: The effect of different diagnostic criteria on the prevalence of dementia. N Engl J Med 337:1667-1774, 1997



Heterogeneity due to data source in OMOP (self-controlled case series) (from OMOP presentations <http://omop.org/2013Symposium>)



Obvious situations for Observational database studies vs. RCT?

- Ethical considerations
- Some safety endpoints
- Rare effects where costs and time would be prohibitive
- Rare diseases: small sample sizes lend themselves to observation vs. randomization
- Study effects of 'off-label' use/ new indications
- To provide historical controls for single-arm studies



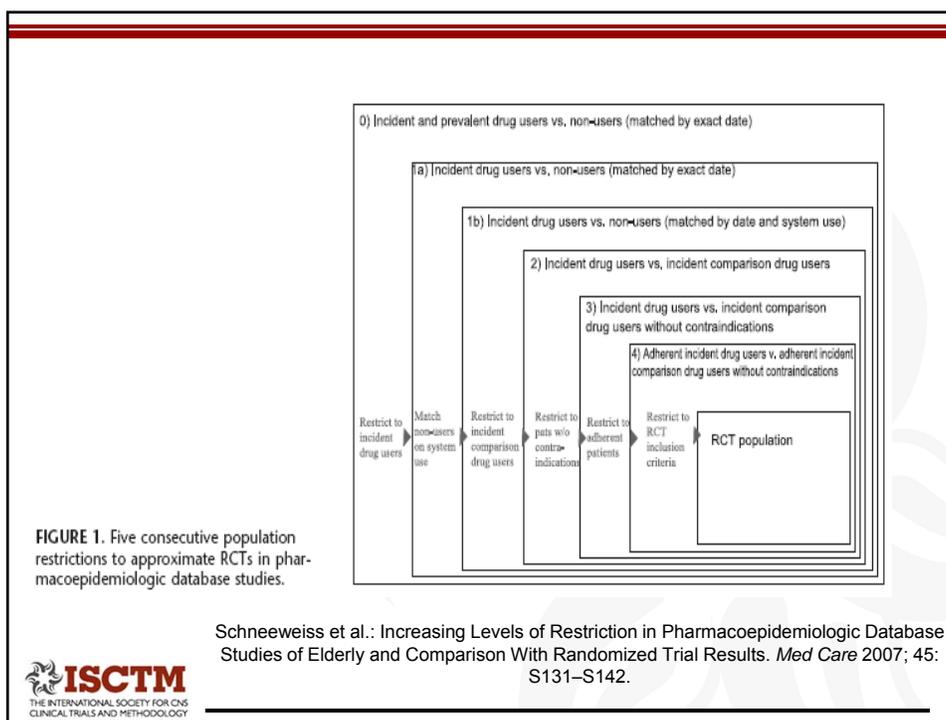
Databases hold a lot of promise

- Large patient populations reflects variability in patients and use
 - Predictive modeling
- Many 'natural experiments' to raise hypotheses
 - **NOTE:** Statistical association \neq causal association
- Capture of 'real world experience'
 - "Adverse events" that arise in current use (some may be benefits)
 - 'Off Label' use can be driven by or lead to insights
- **Poor representation of most benefits**
- Statistical methods to assure comparable patient groups have progressed over last 10 years (mimic RCT)
- Expert clinical involvement can help reduce variability in definitions and interpretation of findings



We are finding that well-designed observational studies approximate the effects of treatment as well as randomized controlled trials on the same topic.

Concato et al: NEJM 2000; 342: 1887-1892.
Benson and Hartz: NEJM 2000; 342:1878-1886.



Cochrane Review

Our results provide little evidence for significant effect estimate differences between observational studies and RCTs, regardless of specific observational study design, heterogeneity, inclusion of pharmacological studies, or use of propensity score adjustment. Factors other than study design *per se* need to be considered when exploring reasons for a lack of agreement between results of RCTs and observational studies.

Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. Cochrane Database of Systematic Reviews 2014, Issue 4.



Statistical Approaches from Real World to RCT: Examples

- **Propensity score:** probability of treatment assignment, conditional on baseline covariates
- **Prior event rate ratio** adjustment (Weiner and Tannen)
 - assumes that the hazard ratio of the exposed to unexposed for a specific outcome before the start of the study reflects the combined effect of all confounders (both measured and unmeasured) independent of any influence of treatment.



Research and Regulatory Environmental Forces

- Personalized medicine/Predictive analytics
 - Population mean giving way to individual prediction
- Real-world effectiveness
- Compressing research cycles in critical areas
- Coverage with Evidence Development



Hybrid Designs: The Intermediate Step

- Pragmatic trials/Large Simple studies
 - Eg, ARTIST
- Bedside Randomization/Randomizing into the Database
- Adaptive design using observational data to develop predictive model of response



Perception/Anticipation/Confirmation Bias

“I wouldn’t have seen it
if I hadn’t believed it”

Attributed to Yogi Berra, Marshall McLuhan

Our ability to detect depends on....

- Background rate of event vs number of patients seen
 - Larger differences more likely to be detected with smaller sample sizes
- How striking the event is
- Singular vs. cluster of symptoms
- Appearance in time/space/proximity to others
- Cognitive bias (are we looking for it?)

Summary

- Availability of large, patient level databases is attractive and will contribute to identification of new uses by raising possibilities
- Databases can make further study more efficient
- The compound must be on the market and captured in the database in order to be studied
- Improved statistical methods have reduced the gap between database studies and RCTs
- Hybrid designs hold promise
- Stay tuned