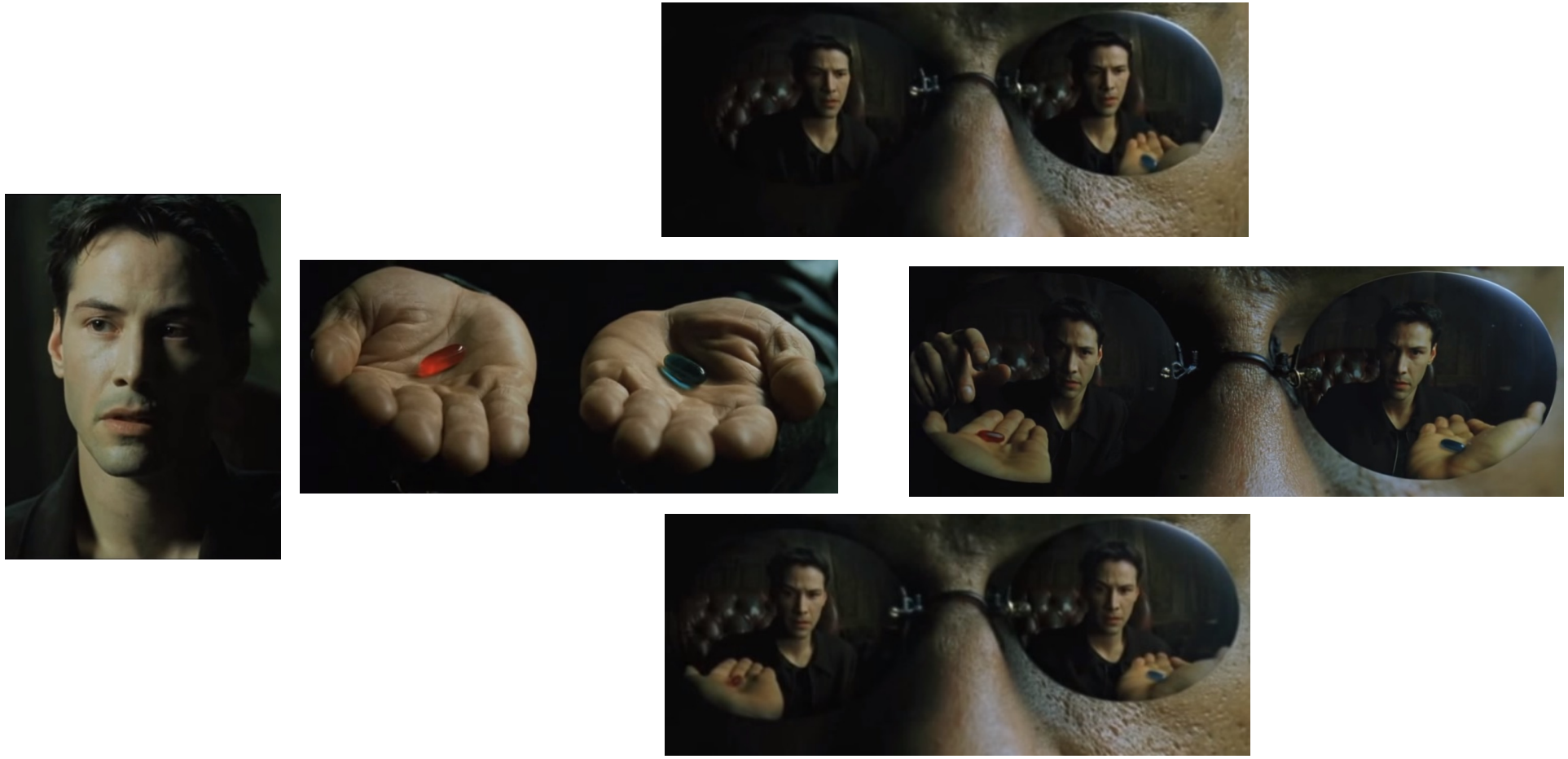
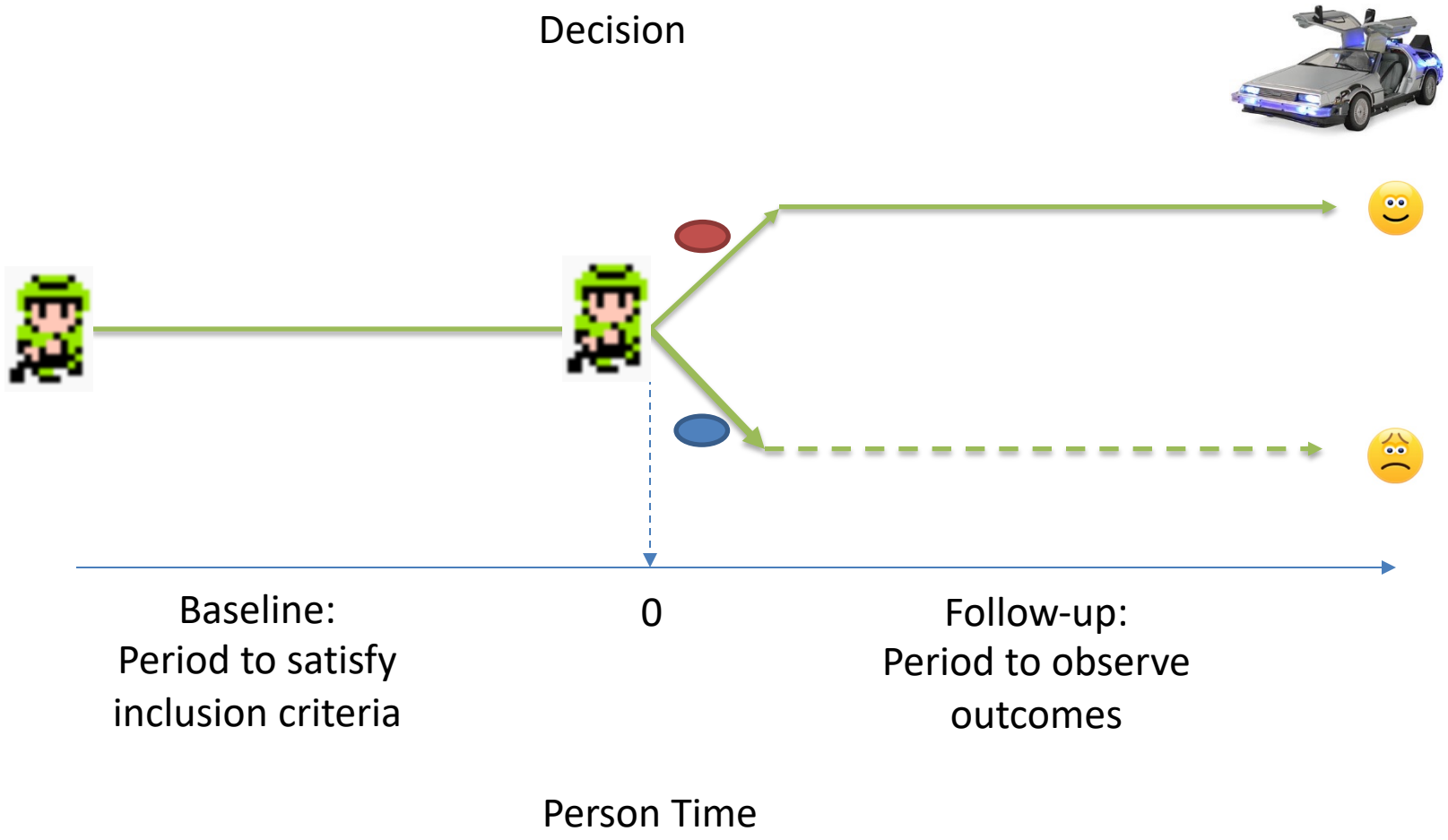
Causal effect estimation

Martijn Schuemie

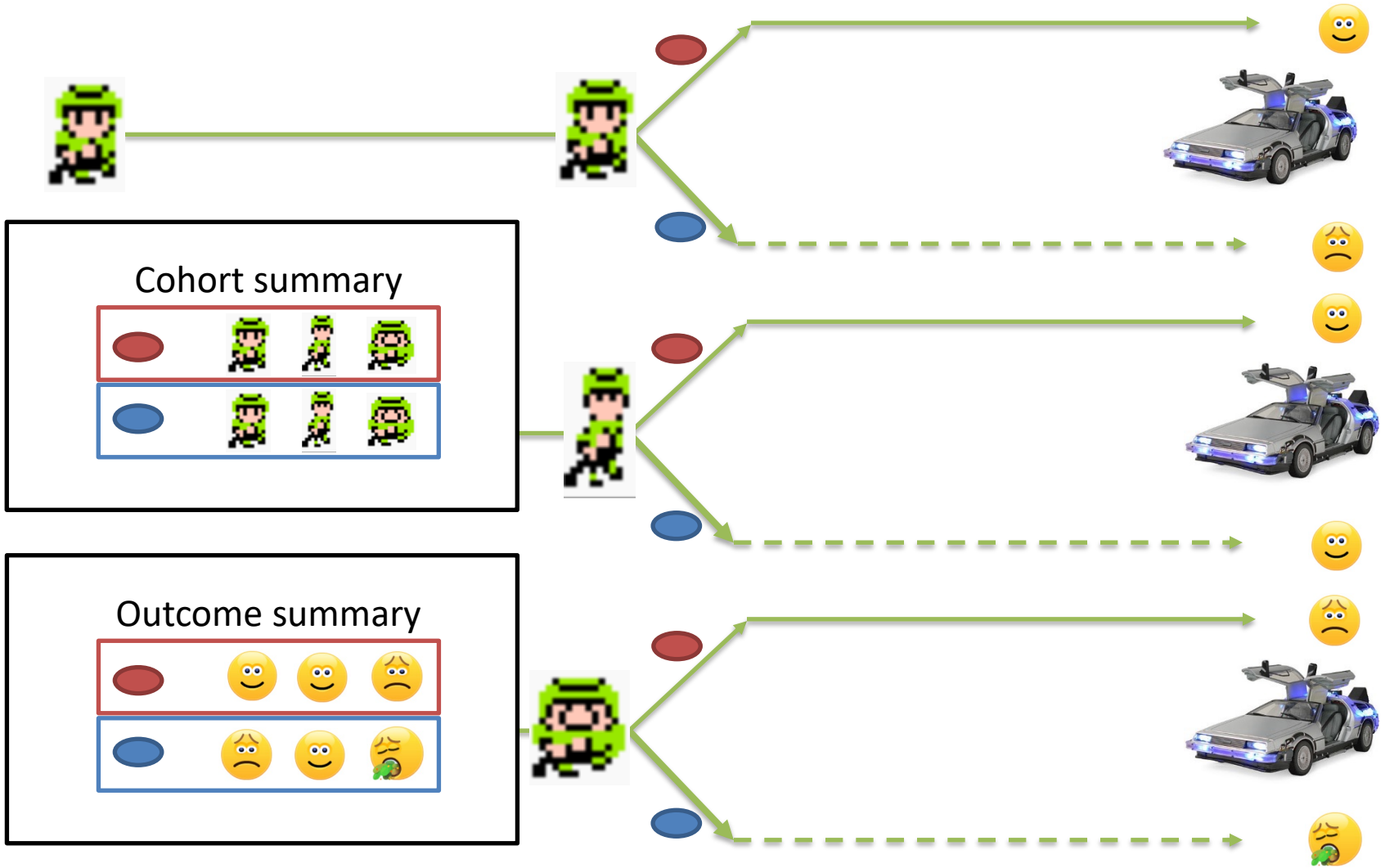
A pop culture mash-up to explain counterfactual reasoning…



Counterfactual reasoning for one person



Counterfactual reasoning for a population

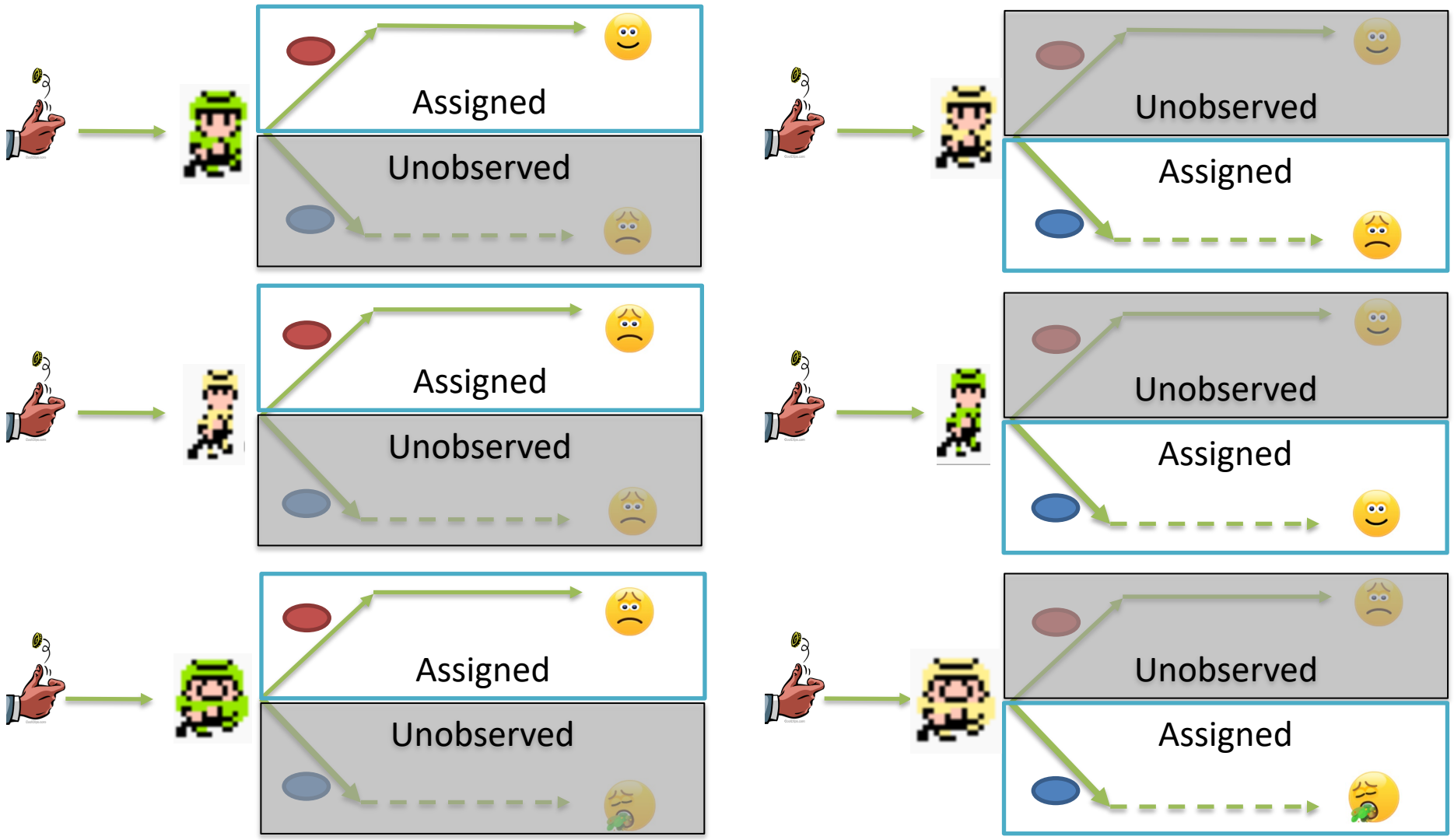


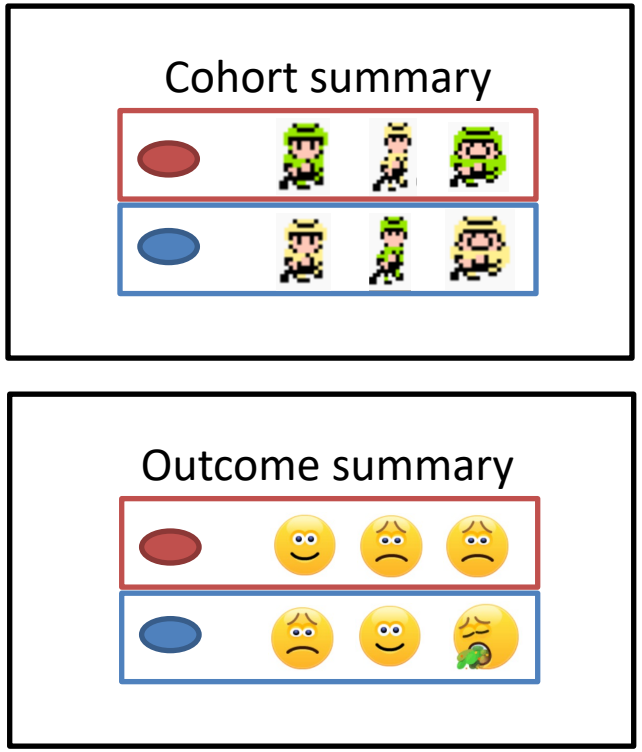
Alas, we don’t have a Delorean…

- What is our *next* best approximation?

- Randomized trial

Randomized treatment assignment to approximate counterfactual outcomes





- Randomization allows for assumption that persons assigned to target cohort are exchangeable at baseline with persons assigned to comparator cohort

Alas, we can’t randomize…

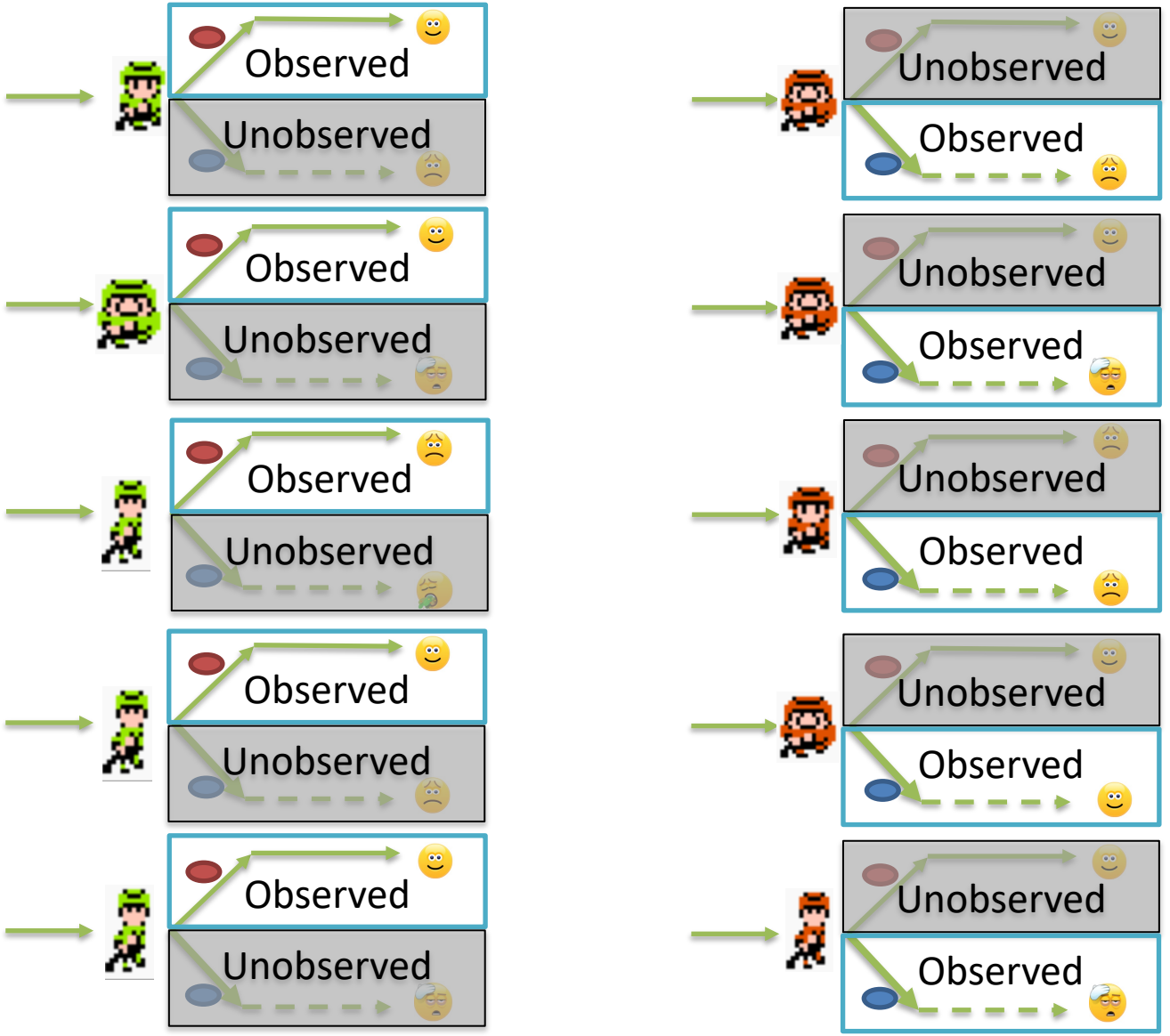
- What is our *next, next* best approximation?

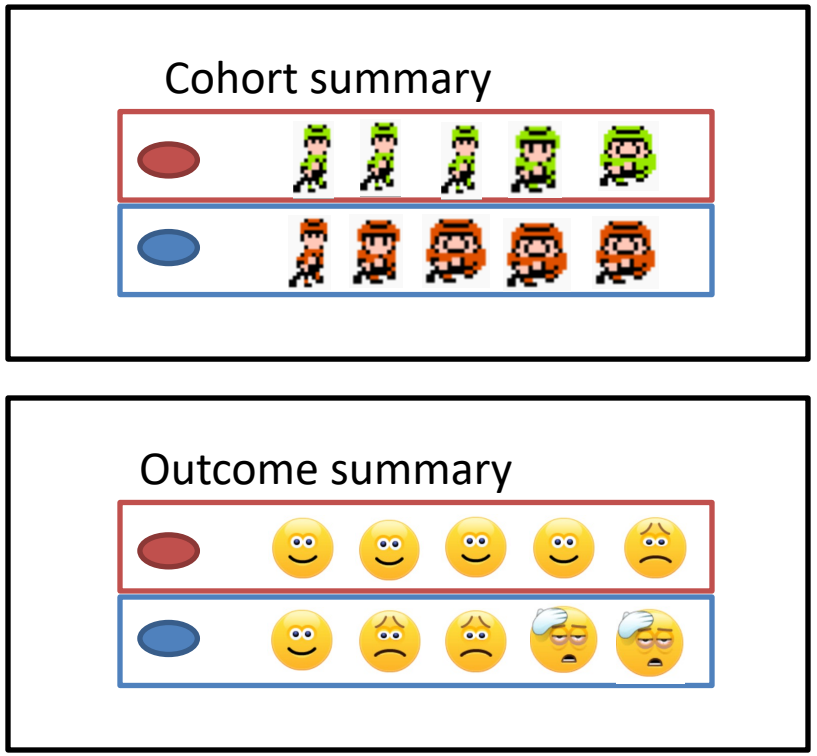
- Observational study:
  - Comparative cohort design: Between persons who made different choices

  OR

  - Self-controlled designs: Within persons during time periods with different exposure status

An observational comparative cohort design to approximate counterfactual outcomes





- Exchangeability assumption may be violated if there is reason for treatment choice...and there often is

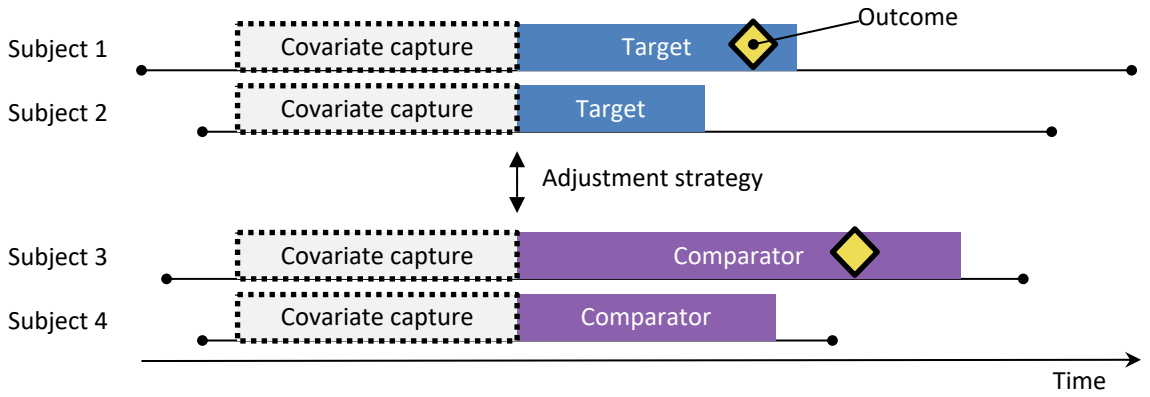
Propensity score introduction

- Propensity score = probability of belonging to the target cohort vs. the comparator cohort, given the baseline covariates
- Pr(Z=1|x)
  - Z is treatment assignment
  - x is a set of all covariates at the time of treatment assignment
- Propensity score can be used as a ‘balancing score’: if the two cohorts have similar propensity score distribution, then the distribution of covariates should be the similar (need to perform diagnostic to check)

Rubin Biometrika 1983

Large-scale propensity scores

- Traditional: select handful of variables to use as predictors of treatment assignment
- OHDSI approach: use all data prior to treatment assignment
  - Conditions
  - Drugs
  - Procedures
  - Observations



- **Important**: fully automated, except you must manually remove target and comparator concepts from the covariates!

Methods for confounding adjustment using a propensity score

Regression adjustment	The PS is used as a covariable in an outcome regression model to adjust the as… assun… same… relationship between propensity score and outcome is correctly specified.
Matching	The PS is used to match exposed subjects to unexposed subjects with similar values of the PS. This method assumes that within the matched sample, exposed and unexposed subjects have a similar distribution of baseline characteristics.
Stratification	The PS is used to stratify subjects into (often quintiles or deciles) strata. Treatment effects are estimated separately within each stratum and then combined into an overall estimate of treatment effect. This method assumes that within each stratum, exposed and unexposed subjects have a similar distribution of baseline characteristics.
Inverse Probability Weighting	The PS is used to create w… defined as: E*/PS + … characteristics are similar …

* E: exposure

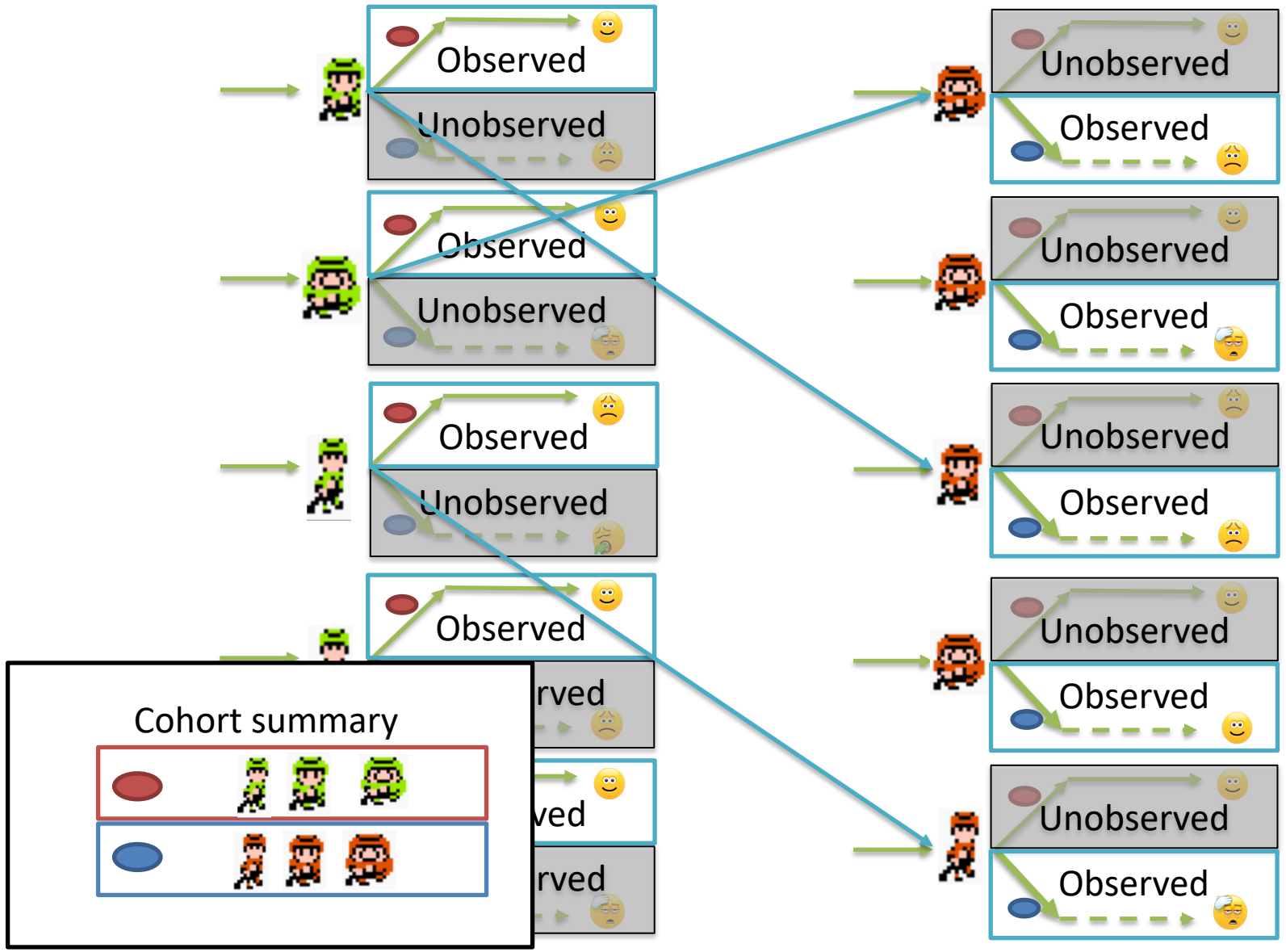
Not generally recommended

Empirical evidence that this doesn't work well

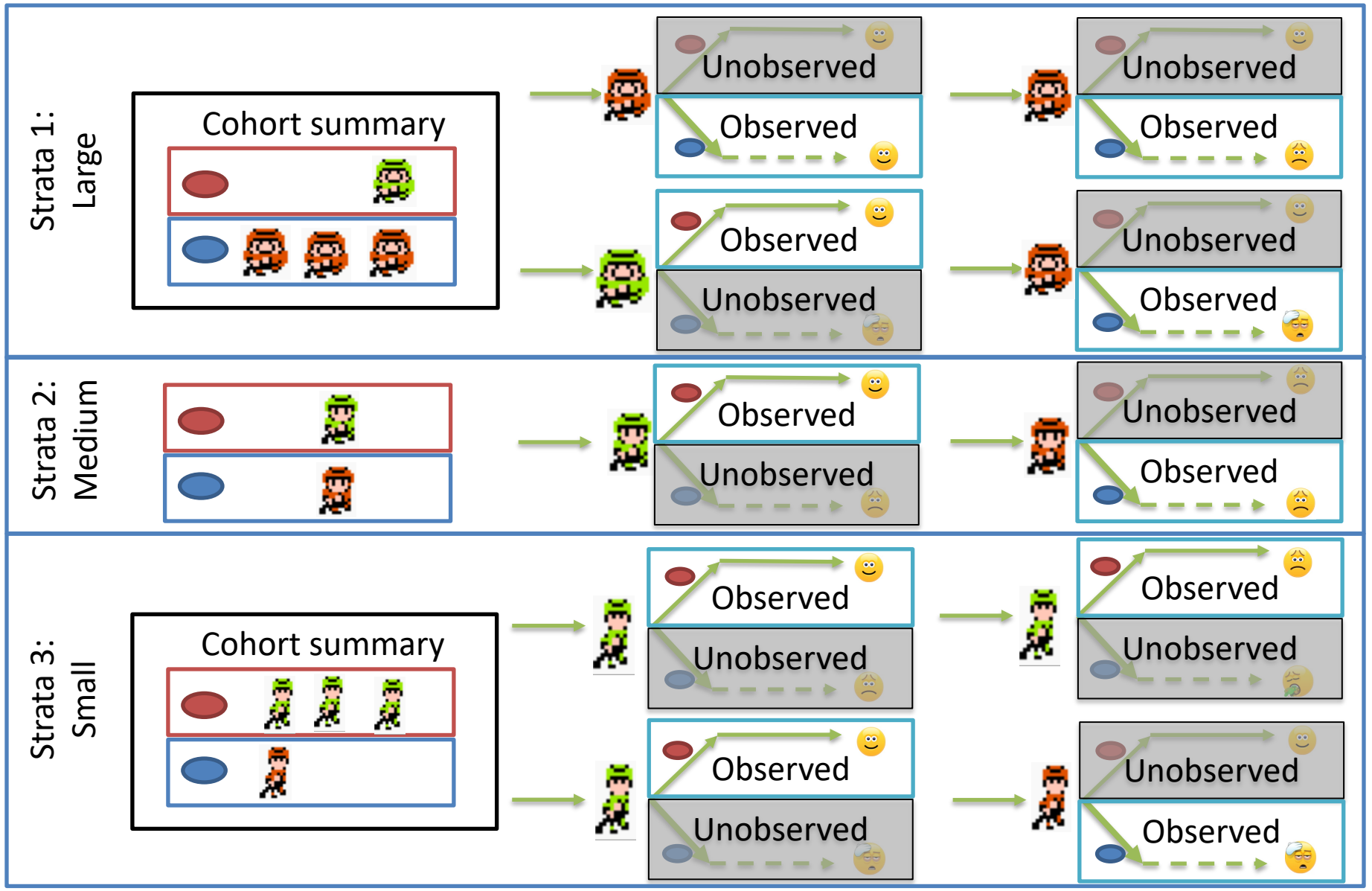
Fully implemented in OHDSI CohortMethod R package

Garbe et al, Eur J Clin Pharmacol 2013, http://www.ncbi.nlm.nih.gov/pubmed/22763756

Matching as a strategy to adjust for baseline covariate imbalance



Stratification as a strategy to adjust for baseline covariate imbalance



The choice of the outcome model defines your research question

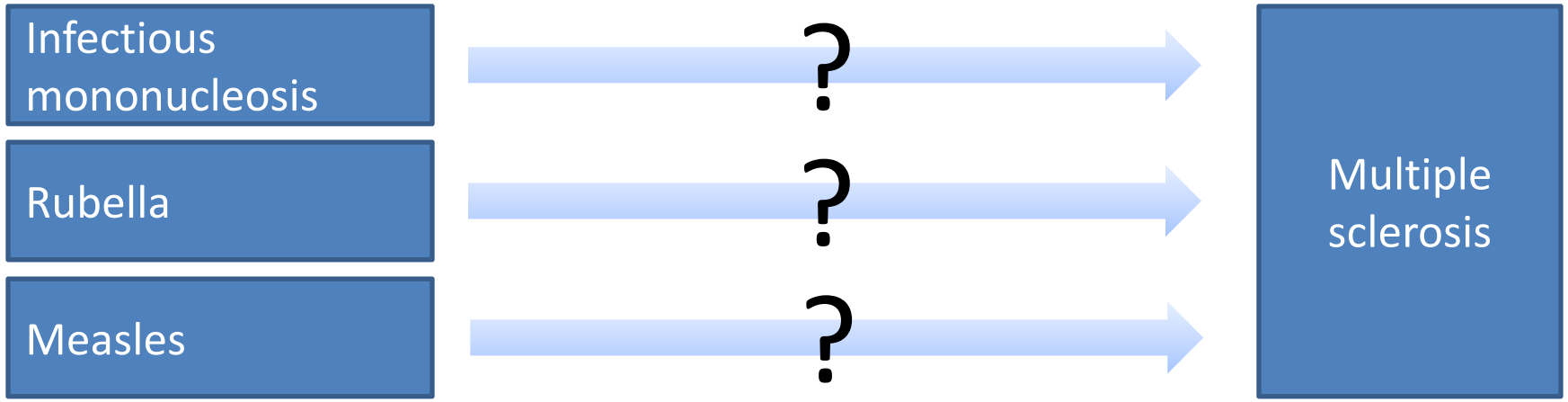
	Logistic regression	Poisson regression	Cox proportional hazards
How the outcome cohort is used	Binary classifier of presence/ absence of outcome during the fixed time-at-risk period	Count the number of occurrences of outcomes during time-at-risk	Compute time-to-event from time-at-risk start until earliest of first occurrence of outcome or time-at-risk end, and track the censoring event (outcome or no outcome)
'Risk' metric	Odds ratio	Rate ratio	Hazard ratio
Key model assumptions	Constant probability in fixed window	Outcomes follow Poisson distribution with constant risk	Proportionality – constant relative hazard

When designing or reviewing a study, ask yourself:

Input parameter	Design choice
Target cohort (T)	
Comparator cohort (C)	
Outcome cohort (O)	
Time-at-risk	
Model specification	

+ negative controls

Examples of negative controls



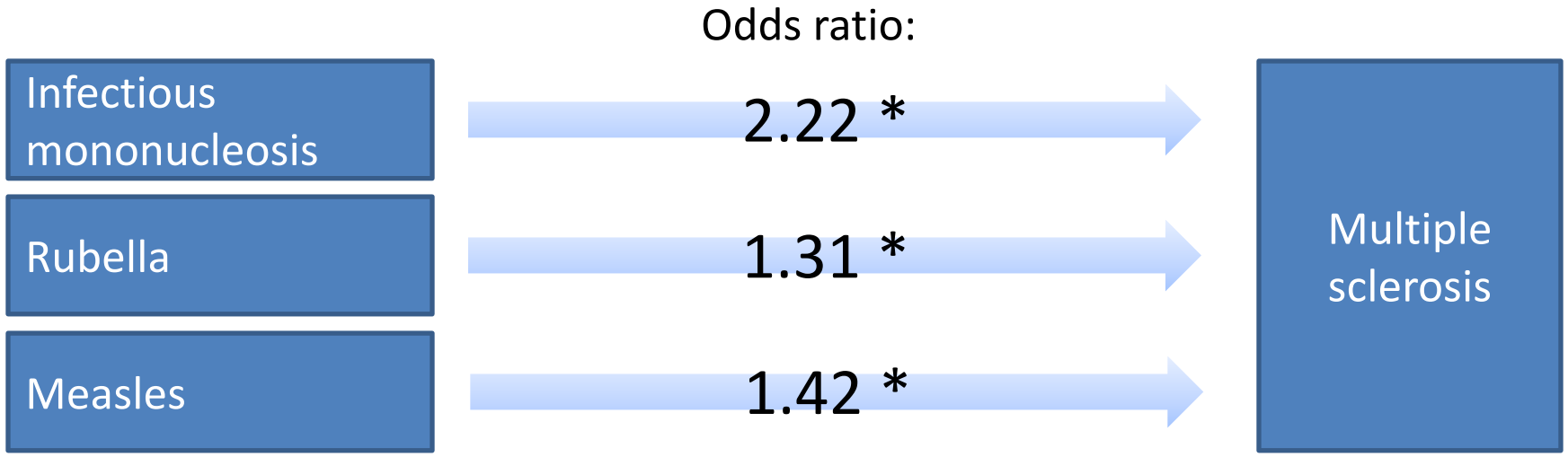
RESEARCH PAPER

Multiple Sclerosis 2008; **14**: 307–313

Selective association of multiple sclerosis with infectious mononucleosis

BM Zaadstra[1,2], AMJ Chorus[1], S van Buuren[1,3], H Kalsbeek[1] and JM van Noort[4]

Example of a negative control



* P < .05

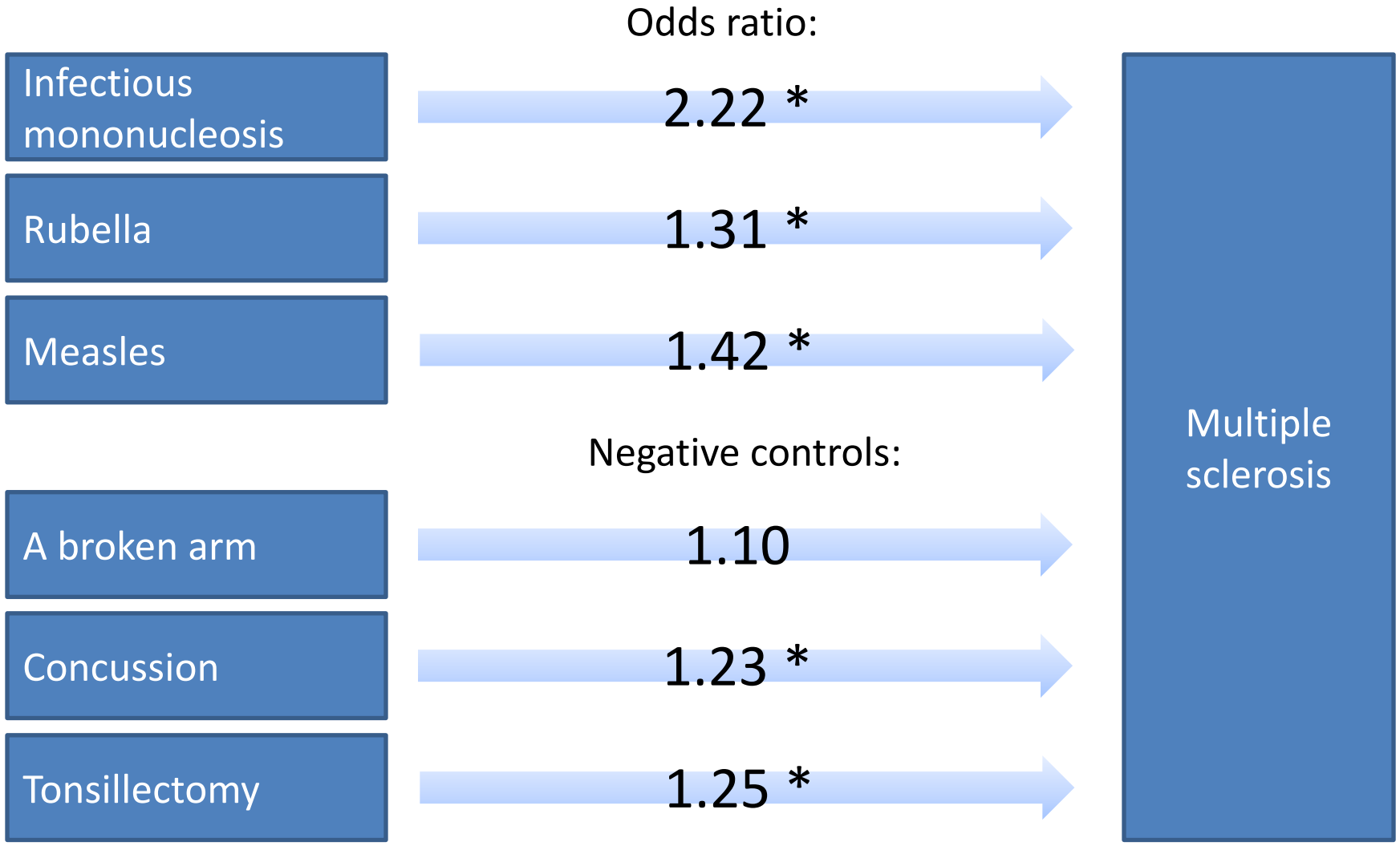
RESEARCH PAPER

Multiple Sclerosis 2008; **14**: 307–313

Selective association of multiple sclerosis with infectious mononucleosis

BM Zaadstra[1,2], AMJ Chorus[1], S van Buuren[1,3], H Kalsbeek[1] and JM van Noort[4]

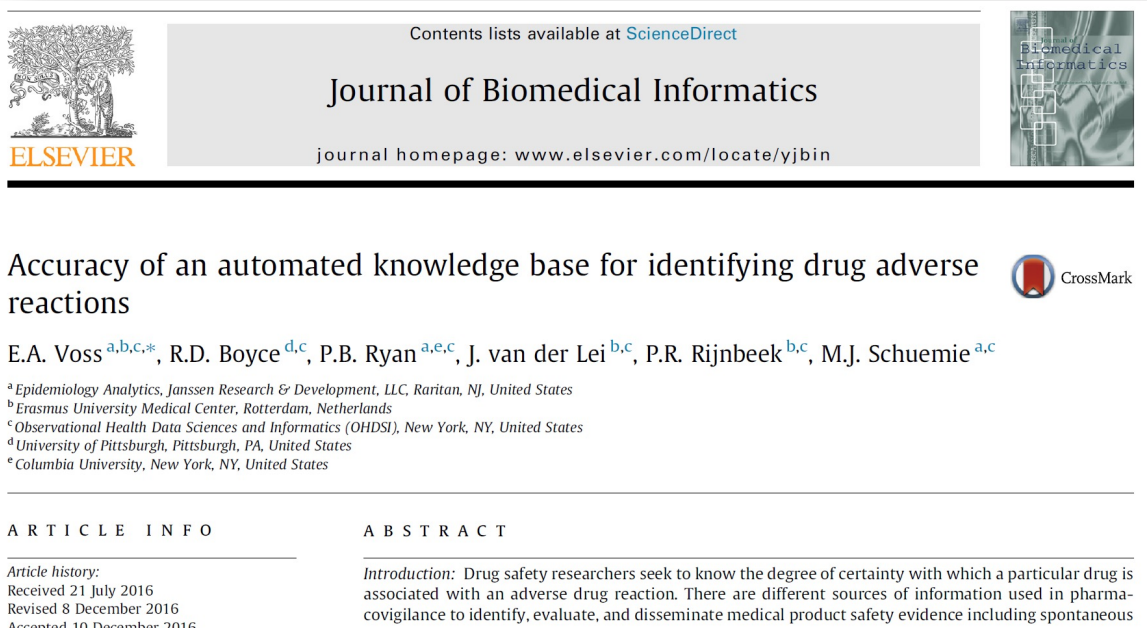
Example of a negative control



* P < .05

Negative controls in a comparative cohort study

- If neither target nor comparator causes the outcome, the hazard ratio / incidence rate ratio / odds ratio should be 1
- Select 50-100 negative control outcomes per study
- ATLAS can help, using information from
  - Product labels
  - Scientific literature
  - Spontaneous reporting



Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Accuracy of an automated knowledge base for identifying drug adverse reactions

E.A. Voss [a,b,c,*], R.D. Boyce [d,c], P.B. Ryan [a,e,c], J. van der Lei [b,c], P.R. Rijnbeek [b,c], M.J. Schuemie [a,c]

[a] Epidemiology Analytics, Janssen Research & Development, LLC, Raritan, NJ, United States
[b] Erasmus University Medical Center, Rotterdam, Netherlands
[c] Observational Health Data Sciences and Informatics (OHDSI), New York, NY, United States
[d] University of Pittsburgh, Pittsburgh, PA, United States
[e] Columbia University, New York, NY, United States

ARTICLE INFO

Article history:
Received 21 July 2016
Revised 8 December 2016

ABSTRACT

Introduction: Drug safety researchers seek to know the degree of certainty with which a particular drug is associated with an adverse drug reaction. There are different sources of information used in pharmacovigilance to identify, evaluate, and disseminate medical product safety evidence including spontaneous