Decomposition models as a framework for thinking about heterogeneity of treatment effects

William Crown, PhD
Chief Scientific Officer

OPTUM Labs®

Overview

1. Randomization solves a lot of problems
2. Proper design is fundamental in observational studies
3. Decomposition of average treatment effects
4. Decomposition with propensity scoring
5. Looks like an RCT. What could go wrong?

Aspects of bias addressed by randomization

- Balances comparison groups on both observed and *unobserved* characteristics
- Greatly simplifies analysis
- Inclusion/exclusion criteria and intensive follow-up in trials introduce issues of generalizability of findings
- Addresses high variability of findings in small trials

RWD vs. RWE (and types of RWE)

The NEW ENGLAND JOURNAL of MEDICINE

SOUNDING BOARD

Real-World Evidence — What Is It and What Can It Tell Us?

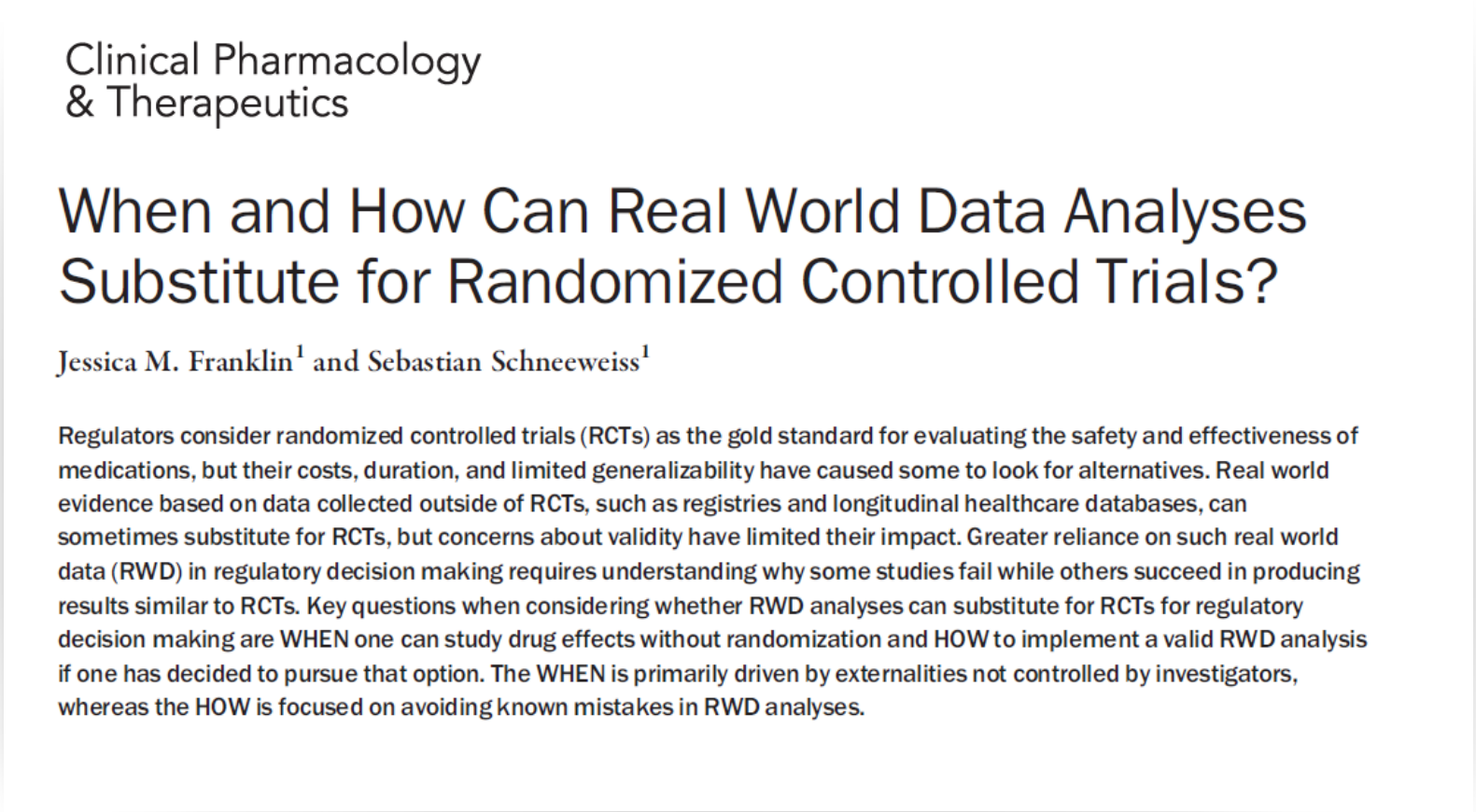
Rachel E. Sherman, M.D., M.P.H., Steven A. Anderson, Ph.D., M.P.P., Gerald J. Dal Pan, M.D., M.H.S., Gerry W. Gray, Ph.D., Thomas Gross, M.D., M.P.H., Nina L. Hunter, Ph.D., Lisa LaVange, Ph.D., Danica Marinac-Dabic, M.D., Ph.D., Peter W. Marks, M.D., Ph.D., Melissa A. Robb, B.S.N., M.S., Jeffrey Shuren, M.D., J.D., Robert Temple, M.D., Janet Woodcock, M.D., Lilly Q. Yue, Ph.D., and Robert M. Califf, M.D.

The term "real-world evidence" is widely used by those who develop medical products or who study, deliver, or pay for health care, but its specific meaning is elusive. We believe it refers to information on health care that is derived from multiple sources outside typical clinical research settings, including electronic health records (EHRs), claims and billing data, product and disease registries, and data gathered through per-

shortage of researchers with adequate methodologic savvy could result in poorly conceived study and analytic designs that generate incorrect or unreliable conclusions. Accordingly, if we are to realize the full promise of such evidence, we must be clear about what it is and how it can be used most effectively, and we must have appropriate expectations about what it can tell us. It is important to distinguish two key dimensions

Sherman R., Anderson S., Dal Pan G., et al. (2016) Real world evidence—what is it and what can it tell us? N Engl J Med; 375(23):2293-2297

Use of RWE vs. RCTs



Clinical Pharmacology & Therapeutics

When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials?

Jessica M. Franklin[1] and Sebastian Schneeweiss[1]

Regulators consider randomized controlled trials (RCTs) as the gold standard for evaluating the safety and effectiveness of medications, but their costs, duration, and limited generalizability have caused some to look for alternatives. Real world evidence based on data collected outside of RCTs, such as registries and longitudinal healthcare databases, can sometimes substitute for RCTs, but concerns about validity have limited their impact. Greater reliance on such real world data (RWD) in regulatory decision making requires understanding why some studies fail while others succeed in producing results similar to RCTs. Key questions when considering whether RWD analyses can substitute for RCTs for regulatory decision making are WHEN one can study drug effects without randomization and HOW to implement a valid RWD analysis if one has decided to pursue that option. The WHEN is primarily driven by externalities not controlled by investigators, whereas the HOW is focused on avoiding known mistakes in RWD analyses.

Franklin J, Schneeweiss S. When and how can real-world data analyses substitute for randomized controlled trials? *Clinical Pharmacology and Therapeutics*. 2017.

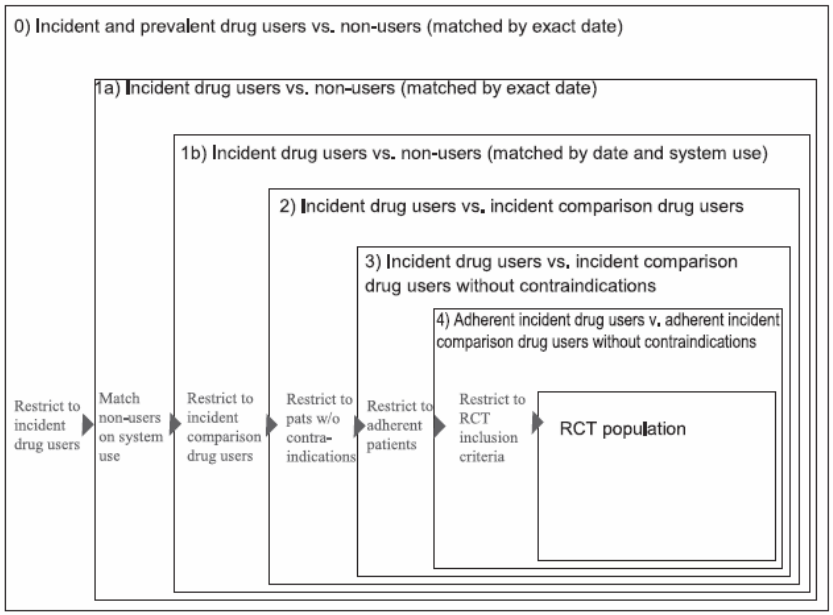
Guideposts that can improve reliability of database study results

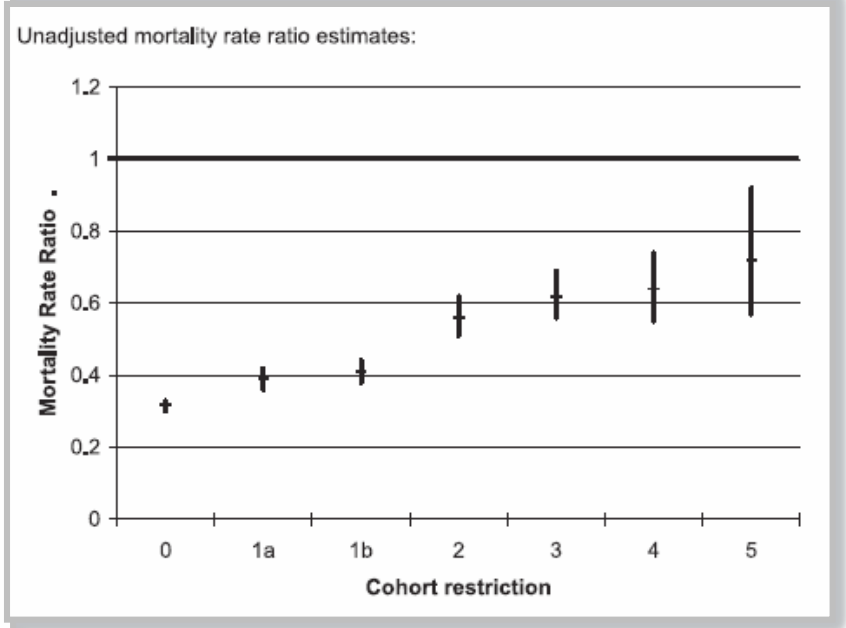
- Active comparator, same treatment modality
- Control for medication adherence
- High-dimensional proxy adjustment
- New users
- Avoid depletion of ‘susceptibles’ and other design flaws

OPTUM Labs®

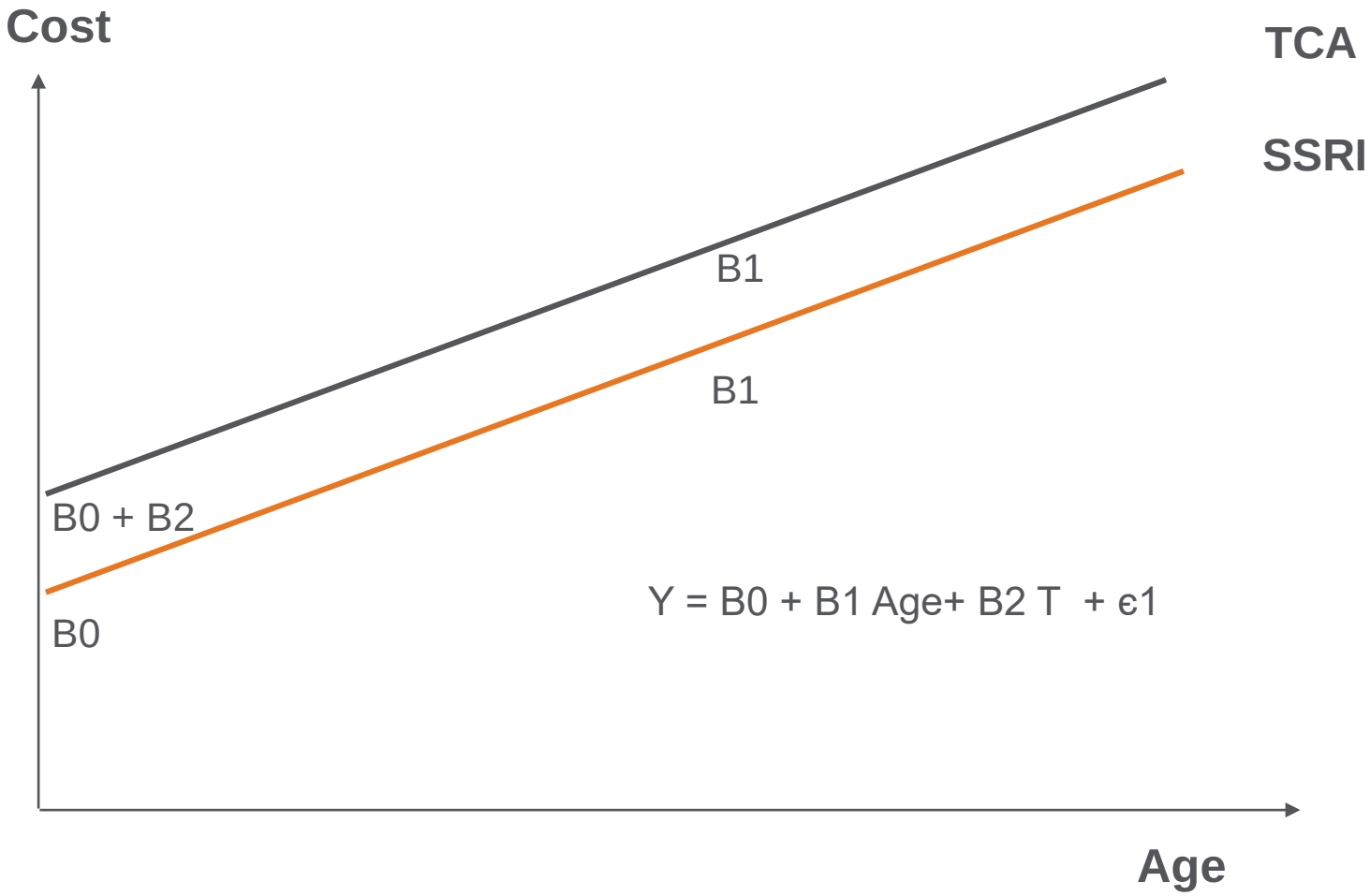
Increasing Levels of Restriction in Pharmacoepidemiologic Database Studies of Elderly and Comparison With Randomized Trial Results

Sebastian Schneeweiss, MD, ScD, Amanda R. Patrick, MS,* Til Stürmer, MD, MPH,* M. Alan Brookhart, PhD,* Jerry Avorn, MD,* Malcolm Maclure, ScD,* Kenneth J. Rothman, DMD, DrPH,† and Robert J. Glynn, PhD, ScD**

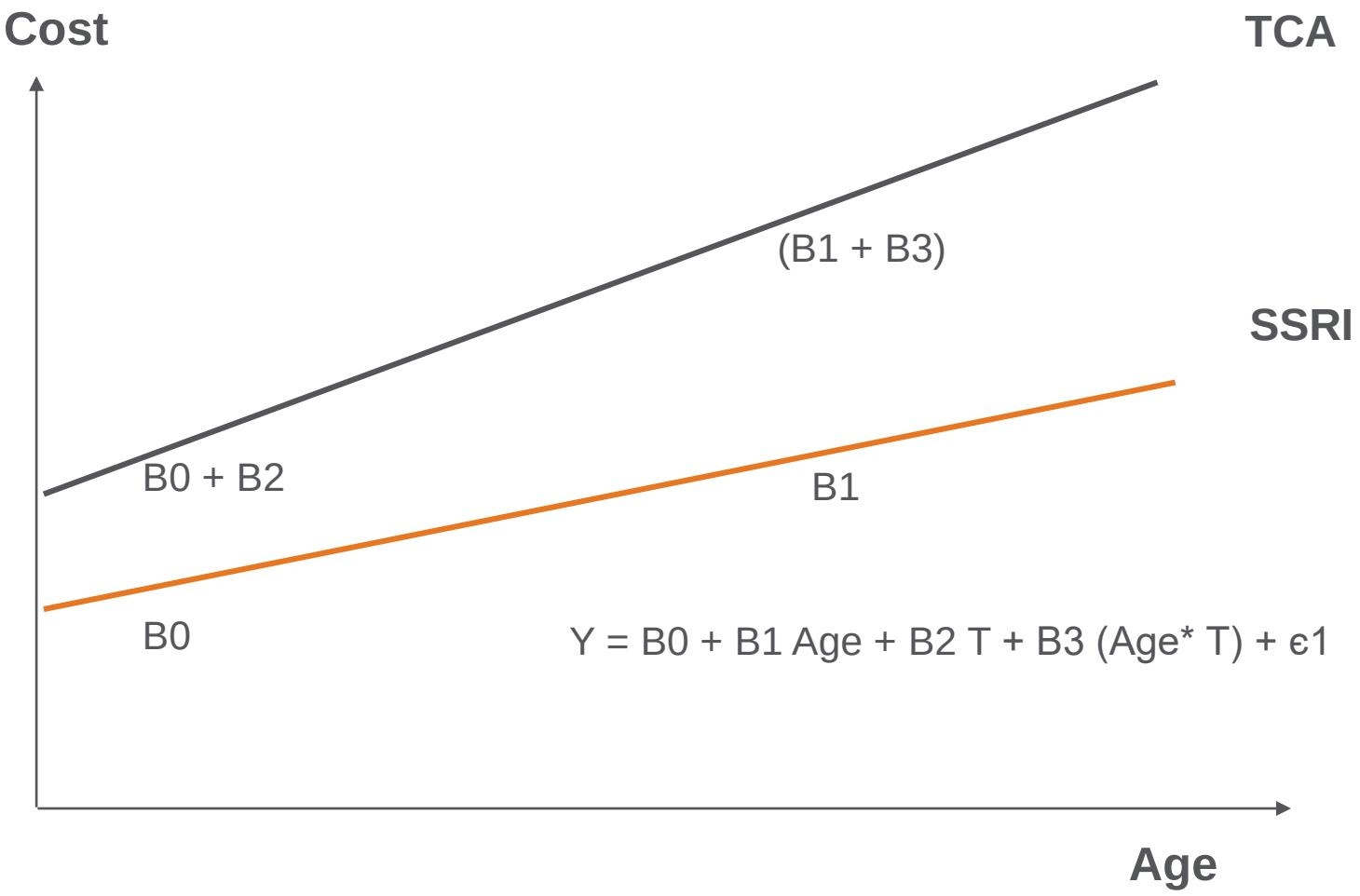




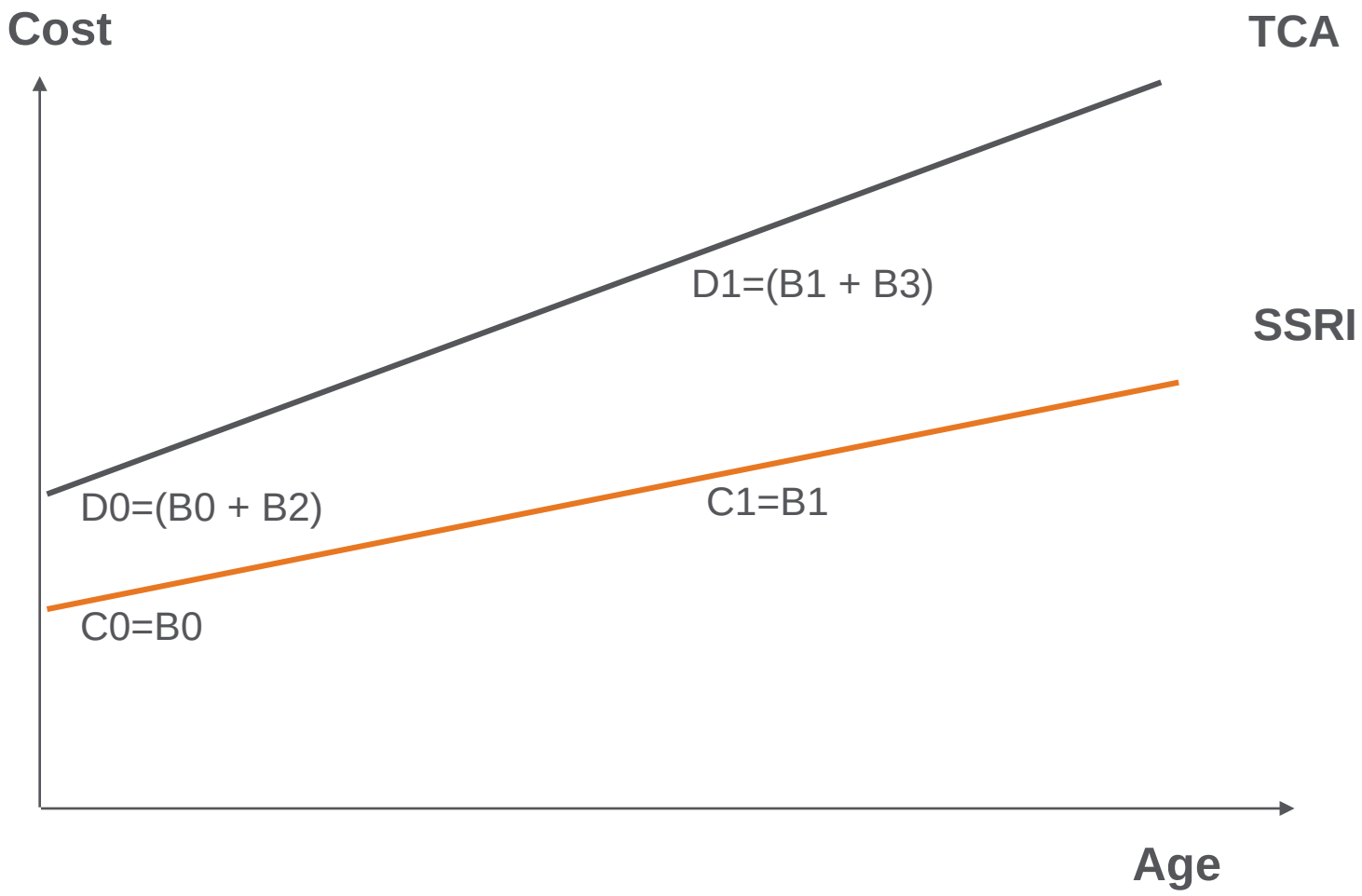
Stand-alone dummy variable



Stand-alone dummy variable and interaction



Simple stratification



Heterogeneity of treatment effects in observational studies

- Heterogeneity of treatment response: in observational studies treatment may interact with covariates resulting in differential patient response.
- Although the covariance between treatment and the covariates is zero in an RCT by design, it is still possible for patients in each treatment group to have heterogeneous treatment responses.

Sample Stratification Versus Interaction Terms

- Sample Stratification solves the multicollinearity problem often associated with interactions, assuming there is sufficient sample size to estimate the separate models

But how can the average treatment effects be reconstructed from the separate equations?

Blinder-Oaxaca Decomposition

$$\overline{Y_{T2}} - \overline{Y_{T1}} = \underbrace{\left[\overline{F(X_{T2}\beta_{T2})} - \overline{F(X_{T1}\beta_{T2})}\right]}_{\text{Composition}} + \underbrace{\left[\overline{F(X_{T1}\beta_{T2})} - \overline{F(X_{T1}\beta_{T1})}\right]}_{\text{Coefficients}}$$

The counterfactual concept:

- Substitute TCA sample through SSRI model (or vice versa)
- Perform many times with subsamples
- Statistical bootstrapping to obtain standard errors

Blinder-Oaxaca decomposition in RCTs

- Blinder and Oaxaca showed that, in observational studies, the difference in expectation between two groups was a function of differences in the sample composition of the groups as well as coefficient effects.
- RCTs, by design, balance sample composition across groups on both observable and unobservable variables. This causes the sample composition terms to drop from the decomposition equation.
- As a result, average treatment effects are equal to a weighted average of the coefficients (heterogeneity of treatment effects).

VALUE IN HEALTH 20 (2017) 1252–1259

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

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

The Use of Decomposition Methods in Real-World Treatment Benefits Evaluation for Patients with Type 2 Diabetes Initiating Different Injectable Therapies: Findings from the INITIATOR Study

Lee Brekke, PhD[1,], Erin Buysman, MS[1], Michael Grabner, PhD[2], Xuehua Ke, PhD[2], Lin Xie, MS[3], Onur Baser, PhD[3,4,5], Wenhui Wei, PhD[6]*

[1]*Optum, Eden Prairie, MN, USA;* [2]*HealthCore, Inc., Wilmington, DE, USA;* [3]*STATinMED Research, Ann Arbor, MI, USA;* [4]*University of Michigan, Ann Arbor, MI, USA;* [5]*School of Economy, Administrative and Social Sciences, MEF University, Istanbul, Turkey;* [6]*Sanofi US, Bridgewater, NJ, USA*

Overall decomposition results – HbA1c

Table 2 – HbA_{1c} changes: Results of overall decomposition analysis at 1 y (N = 2166).

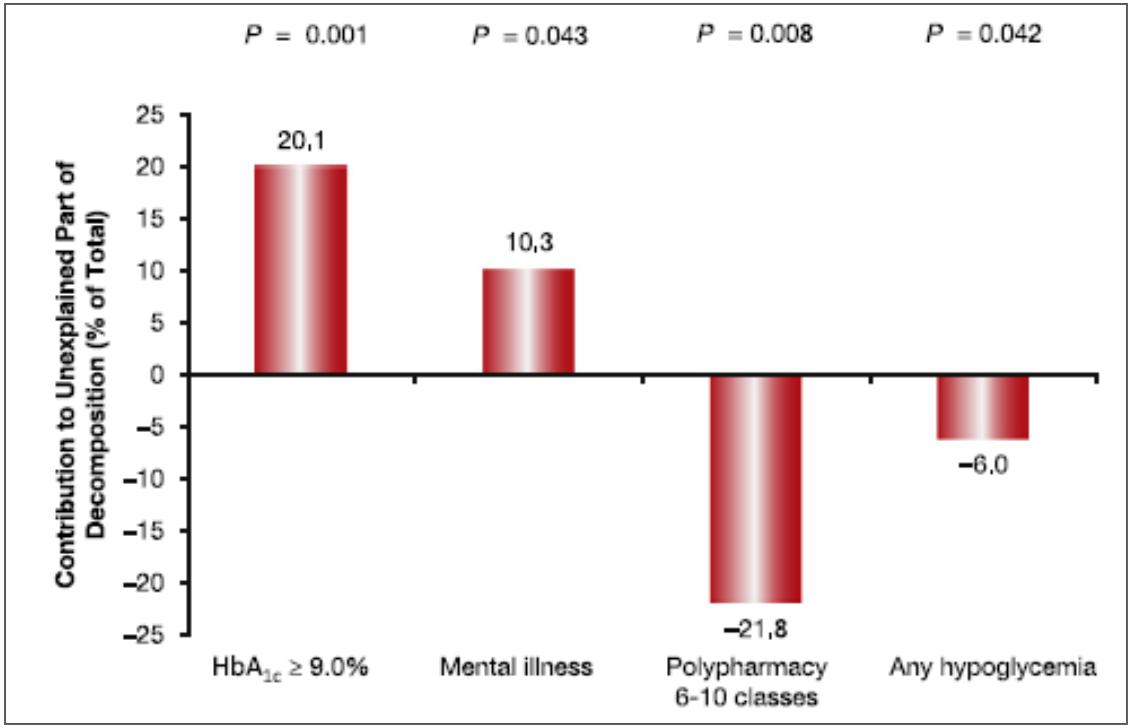
	GLA (n = 1107)	LIRA (n = 1059)	Difference (GLA – LIRA)	Explained difference (attributed to differences in baseline characteristics)	Unexplained difference (attributed to differences in treatment effects)
Mean HbA_{1c} change (%)	−1.387	−0.742	−0.645	−0.730	0.085
P value	<0.001	<0.001	<0.001	<0.001	0.434

GLA, glargine; HbA_{1c}, glycated hemoglobin A_{1c}; LIRA, liraglutide.

Brekke L., Buysman E., Grabner M., Ke X., Xie L., Baser O., Wei W. (2017) The Use of Decomposition Methods in Real-World Treatment Evaluation for Patients with Type-2 Diabetes Initiating Different Injectable Therapies: Findings from the INITIATOR Study. Value in Health 20:1252-1259.

OPTUM Labs®

Components of treatment effects – HbA1c



Brekke L., Buysman E., Grabner M., Ke X., Xie L., Baser O., Wei W. (2017) The Use of Decomposition Methods in Real-World Treatment Evaluation for Patients with Type-2 Diabetes Initiating Different Injectable Therapies: Findings from the INITIATOR Study. Value in Health 20:1252-1259.

OPTUM Labs®

Overall decomposition results – treatment persistence

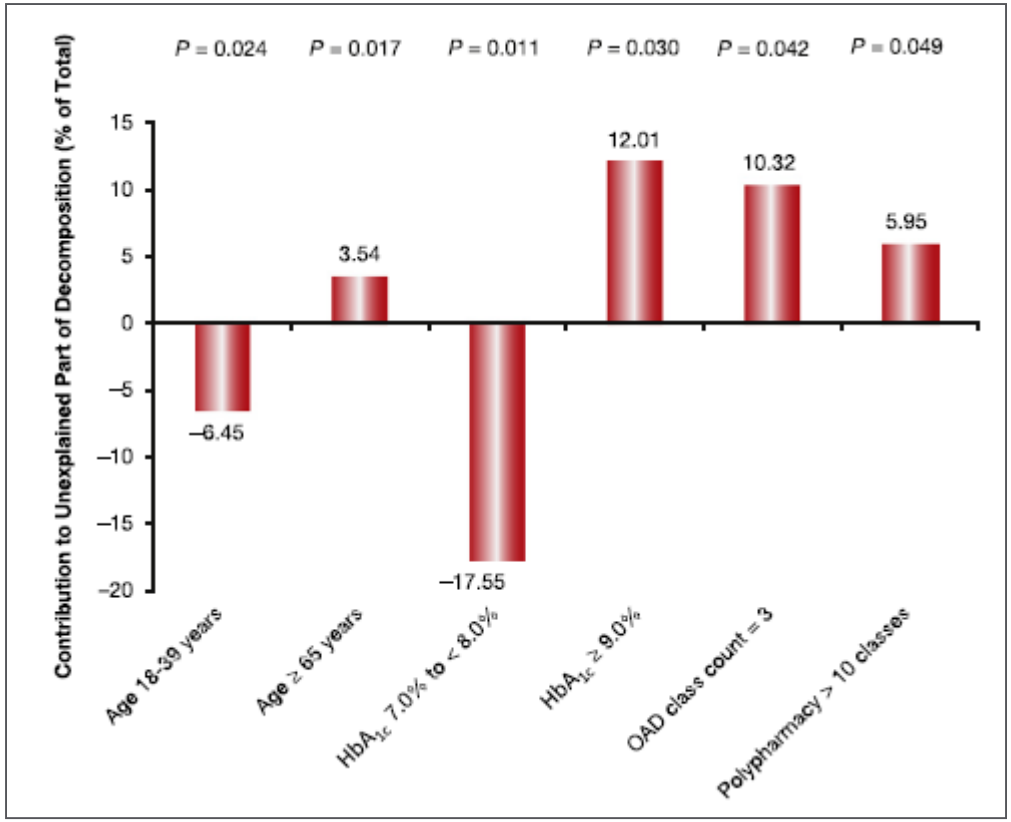
Table 3 – Treatment persistence: Results of overall decomposition analysis at 1 y (N = 3010).

	GLA (n = 1572)	LIRA (n = 1438)	Difference (GLA − LIRA)	Explained difference (attributed to differences in baseline characteristics)	Unexplained difference (attributed to differences in treatment effects)
Mean persistence (%)	64.8	48.7	16.1	−1.8	17.9
P value	<0.001	<0.001	<0.001	0.215	<0.001

GLA, glargine; LIRA, liraglutide.

Brekke L., Buysman E., Grabner M., Ke X., Xie L., Baser O., Wei W. (2017) The Use of Decomposition Methods in Real-World Treatment Evaluation for Patients with Type-2 Diabetes Initiating Different Injectable Therapies: Findings from the INITIATOR Study. Value in Health 20:1252-1259.

Components of treatment effects – treatment persistence



Brekke L., Buysman E., Grabner M., Ke X., Xie L., Baser O., Wei W. (2017) The Use of Decomposition Methods in Real-World Treatment Evaluation for Patients with Type-2 Diabetes Initiating Different Injectable Therapies: Findings from the INITIATOR Study. Value in Health 20:1252-1259.

Blinder-Oaxaca decomposition in observational studies with prior matching

- If patients in comparison groups were first balanced via propensity score or other matching methods, the sample composition terms would also drop from the decomposition equation in observational studies (assuming no omitted variables, etc).
- As with RCTs, the average treatment effect would be a weighted average of the heterogeneity of patient response within the groups.
- For otherwise similar data measurement, the difference between results from an observational study and an RCT would result from omitted variables, measurement error, or other factors that might introduce bias in the coefficient estimates in the observational setting.

What can go wrong

- Estimation of treatment effects within subgroups has been shown to lead to high false-positive and false-negative estimates using RCT data (Brookes et al., 2001).
- Results can be different when the reference group is changed (Fortin, Lemieux, Firpo, 2010).
- Bias can be introduced by omitted variables, measurement error, and other issues common in observational analyses.

Summary

- Decomposition methods have been developed in the labor economics literature to estimate gender and racial discrimination in wages.
- The methods decompose total variation between groups into variation that is due to differences in sample characteristics and variation that is due to the relationships of each group's characteristics to their outcomes.
- Randomization is intended to balance the characteristics of comparison cohorts on both observed and unobserved variables.
- Application of decomposition methods to RCT data leads to the conclusion that average treatment effects are a weighted average of structural coefficients within and across each arm of the trial.
- Comparison of RCT and observational studies within the decomposition analysis framework is helpful for identifying the conditions under which estimates of treatment effect would be expected to be similar for the two types of designs.
- Highlights the importance of variable completeness and measurement quality to minimize bias from residual confounding in observational studies.

References

Blinder, A. S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates." The Journal of Human Resources **8**(4): 436-455.

Brekke L., Buysman E., Grabner M., Ke X., Xie L., Baser O., Wei W. (2017) The Use of Decomposition Methods in Real-World Treatment Evaluation for Patients with Type-2 Diabetes Initiating Different Injectable Therapies: Findings from the INITIATOR Study. Value in Health 20:1252-1259.

Brookes ST, Whitley E, Peters TJ, Mulheran PA, Eggar M, Davy Smith G. Subgroup Analyses in Randomized Controlled Trials: Quantifying the Risks of False-Positives and False-Negatives. Health Technology Assessment 2001; 5(33).

Cook B., McGuire T., Zaslavsky A. (2012) .Measuring Racial/Ethnic Disparities in Health Care: Methods and Practical Issues. Health Services Research 37(3 pt 2):1232-1254.

Crown, W. H. (2010). "There's a reason they call them dummy variables: a note on the use of structural equation techniques in comparative effectiveness research." PharmacoEconomics **28**(10): 947-955.

Fortin N., Lemieux T., Firpo S. Decomposition Methods in Economics. Handbook of Labor Economics (2010). Elsevier Press.

Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labor Markets." International Economic Review **14**(3): 693-709.

Oaxaca, R. L. and M. R. Ransom (1994). "On discrimination and the decomposition of wage differentials." Journal of Econometrics **61**(1): 5-21.

Sherman R., Anderson S., Dal Pan G., et al. (2016) Real world evidence—what is it and what can it tell us? N Engl J Med; 375(23):2293-2297.

Thank you.

Contact information:

William H. Crown, CSO OptumLabs

William.crown@optum.com

OPTUM Labs®