

Beyond Regression Analysis

Using Marginal Structural Models to Estimate Program Effects

Joel M. Moskowitz, PhD, Director

Diana D. McDonnell, PhD, Research Epidemiologist

Gene Kazinets, PhD, Statistician

**Center for Family & Community Health
UC Berkeley PRC**



School of
Public Health

UNIVERSITY OF CALIFORNIA, BERKELEY

Funding provided by the Centers for Disease Control and Prevention, Cooperative Agreement #U48/DP000033. Contents are solely the responsibility of the authors and do not necessarily represent official views of the CDC. We would like to thank the Korean American Community Advisory Board (KCAB) for input on this study.



Ideal Program Evaluation



- **Assess Causal Effects**
 - To gain insight into mechanisms
 - Identify key actions that can modify health
- ➔ Obtain with **Marginal Estimates**
 - Mean change in **population** outcome due to program exposure



Randomized Controlled Trials

- If perfect, provide marginal effects
- Almost never perfect
 - Cannot always do them
 - Constrained by ethics, practicality, cost, time
 - Potential complications
 - Empirical confounding
 - Missing data
 - Incomplete compliance with program
 - Time varying covariates (in causal pathway)
 - Multiple treatments over time

→ Quasi-Experiment or Observational Study



Marginal Structural Models

- MSM is way to model data that estimates the **population**-wide **causal** effect of a program on the outcome
- **Weight data** to make program & comparisons groups equivalent on all variables that affect the outcome *before* estimating program effects
 - Similar to age adjusted rates & survey weighting



MSM: 3 Steps

- Mechanics
 1. Model the **treatment mechanism** (probability of exposure to program)
 - Function of covariates – best fit
 2. Use treatment model to **calculate weights** for each person
 - Inverse probabilities
 3. Do **weighted regression**
 - Outcome as a function of the treatment



MSM Assumptions

- **Same as conventional regression**
 - No unmeasured confounding
 - Time-order is respected explicitly
 - Exposure to program & covariates must *precede* outcome
 - Experimental treatment assignment (ETA)
 - Probability of exposure to the program for any covariate in the treatment model cannot be all 0 or all 1



Example: Quitting is Winning

- Community-based participatory research
 - Korean Community Advisory Board (KCAB)
 - Prioritized smoking cessation in 2001
 - Selected intervention & program
 - Internet program
 - High access
- Adapted existing program
 - Cognitive-behavioral
 - Internet-based
 - Placebo: Printed booklet





Internet Program Website

Center for Family and Community Health | University of California, Berkeley

Health is Strength Quitting is Winning

한국어 English

Community Links

Online Support Group
Gallery of Public Pledge

Reference Tools

Online Quit Program
Nicotine Dependency Test
Bibliography
Help

You Are Here :

1 About This System

2 Am I ready to Quit?

3 Withdrawal Symptoms

4 Preparing to Quit

5 Maintaining My Quit

Welcome to The Quitting is Winning Program

Whether you have already quit smoking or are thinking about quitting smoking, this web program and Support Community will give you the help you need to succeed!

Assess your level of nicotine dependence, talk to new and experienced quitters in our Support Group, find a quitting buddy, or create a personalized quit program that will increase your chances of success.

You do not have to purchase any products to use this free program. Please remember that this program is not to replace the advice of a healthcare professional.

Click on the statement that best describes you :

- I'm Ready to Quit
- I'm Planning on Quitting
- I've Already Quit
- I Don't Intend to Quit
- Help! I've Slipped
- Someone Close to Me Smokes

My Glove Compartment

Welcome to your Online Quit Program. As you proceed your personal information will be stored [HERE](#).

Quote of the Day

If you want to take your mission in life to the next level, if you're stuck and you don't know how to rise, don't look outside yourself. Look inside. Don't let your fears keep you mired in the crowd. Abolish your fears and raise your commitment level to the point of no return, and I guarantee you that the Champion Within will burst forth to propel you toward



Internet Program Website

가족과 지역사회 건강 센터 | UC 버클리

건강은 힘이다 금연으로 승리하자

한국어 English

커뮤니티 링크

온라인 써포트그룹
공개 서약 갤러리

참고 도구

온라인 금연 프로그램
니코틴 의존도 테스트
참고 자료
도움말

현재 위치

- 1 본 시스템에 관하여
- 2 나는 금연준비가 되어 있는가
- 3 금단증상
- 4 금연준비
- 5 금연유지

금연으로 성공하자 프로그램에 오신 것을 환영합니다.

이미 금연을 하고 계시거나, 아니면 현재 금연을 고려하고 계시면 상관없이 저희 웹프로그램과 써포트 커뮤니티가 성공에 필요한 도움을 드릴 것입니다!

자신의 니코틴 의존도를 측정하고, 저희 써포트그룹의 새 금연자 또는 금연 경험자들과 대화를 하시거나, 성공률을 높여 줄 개인 금연프로그램을 짜십시오.

본 프로그램은 무료이며 제품을 구입하지 않아도 이용하실 수 있습니다. 단, 본 프로그램이 의료진의 조언을 대신하지는 않는다는 것을 기억해 주십시오.

자신을 가장 잘 표현한 문장을 클릭하십시오.

- 나는 금연준비가 되었습니다.
- 나는 금연을 계획 중입니다.
- 나는 이미 금연을 시작했습니다.
- 나는 금연할 의도가 없습니다.
- 도와 주세요! 잠깐 실수했습니다.
- 나와 가까운 사람이 담배를 피웁니다.

홈 | 개인정보 보호정책 | 연락처 | 참여금 |

금연 프로그램은 교육용 유일한 목적으로 하고 있으며 가정의나 다른 의료제공자를 대신하여 상담이나 치료를 해 줄 수는 없습니다.

나의 클럽 컴파트먼트

온라인 금연 프로그램에 오신 것을 환영합니다
프로그램에 참여해 나가는 동안 당신의 개인정보가 **여기** 에 저장됩니다

오늘의 인용구

어디에 나동그라졌는지를 볼 것이 아니라 어디서부터 미끄러지게 됐는지를 보라.
- 아프리카 금연

[로그아웃](#)



Research Overview

- Randomized Controlled Trial
 - 2 experimental conditions
 - Internet program vs. Booklet program (placebo control)
 - Pretest & multiple (10) post-tests
- Sample: 708 Korean American men

Study Research Objectives



- Compare effectiveness of Internet-based, cognitive/behavioral, self-help cessation program with similar Booklet program



Measures

- Dependent Variable: Smoking status
 - 1 = **quit** (smoked **0** cigarette in past 7 days)
 - 0 = **not quit** (≥ 1 cigarettes in past 7 days)
- Covariates
 - Demographics
 - Korean language preference
 - Payment schedule: Installments vs. End
 - Nicotine dependence score (FTND)

Analysis 1: Intent-to-Treat

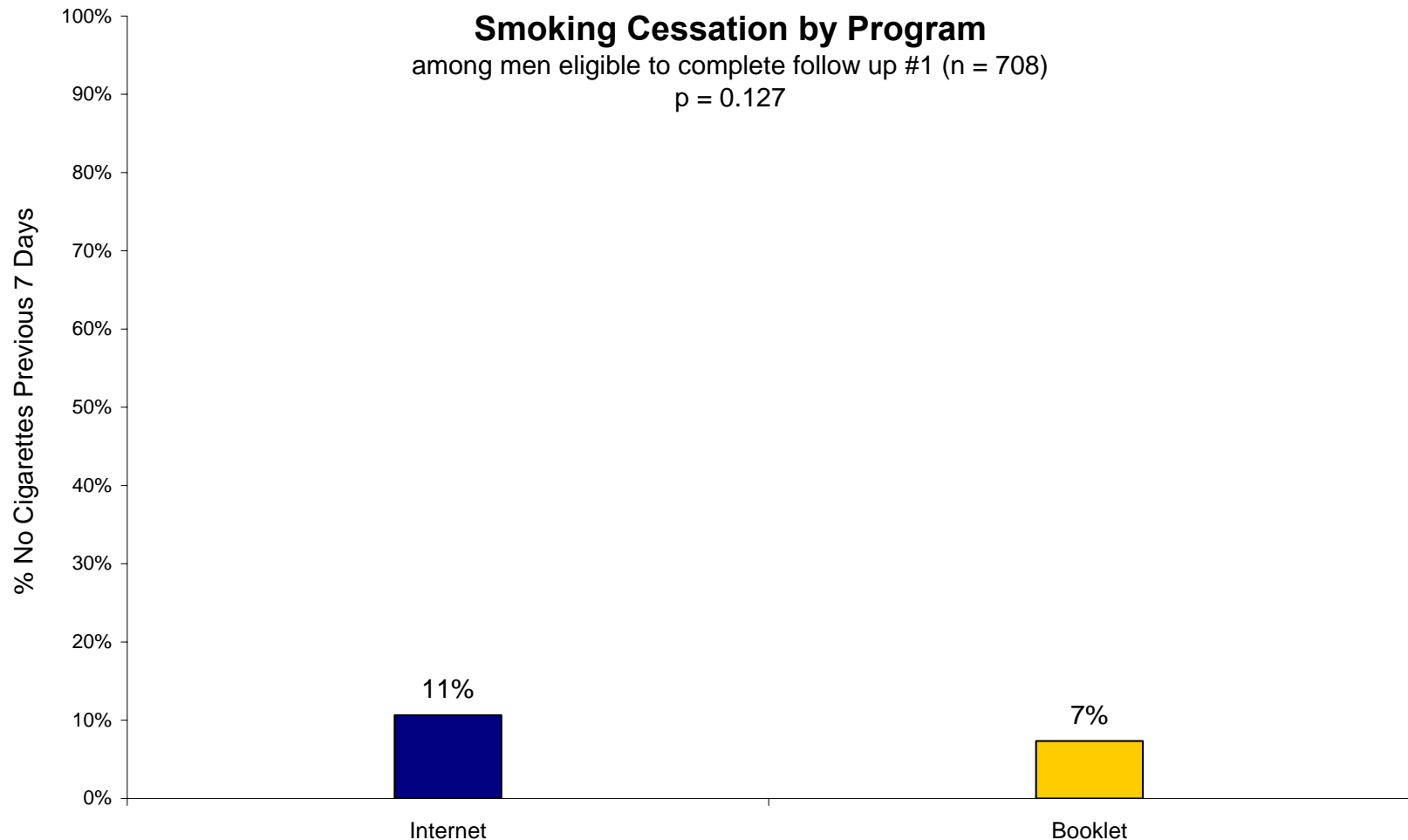


- Internet vs. Booklet
- Based on randomization
- Missing observations code as *not quit*

Quit Smoking



- Trend toward more quitting among internet program.





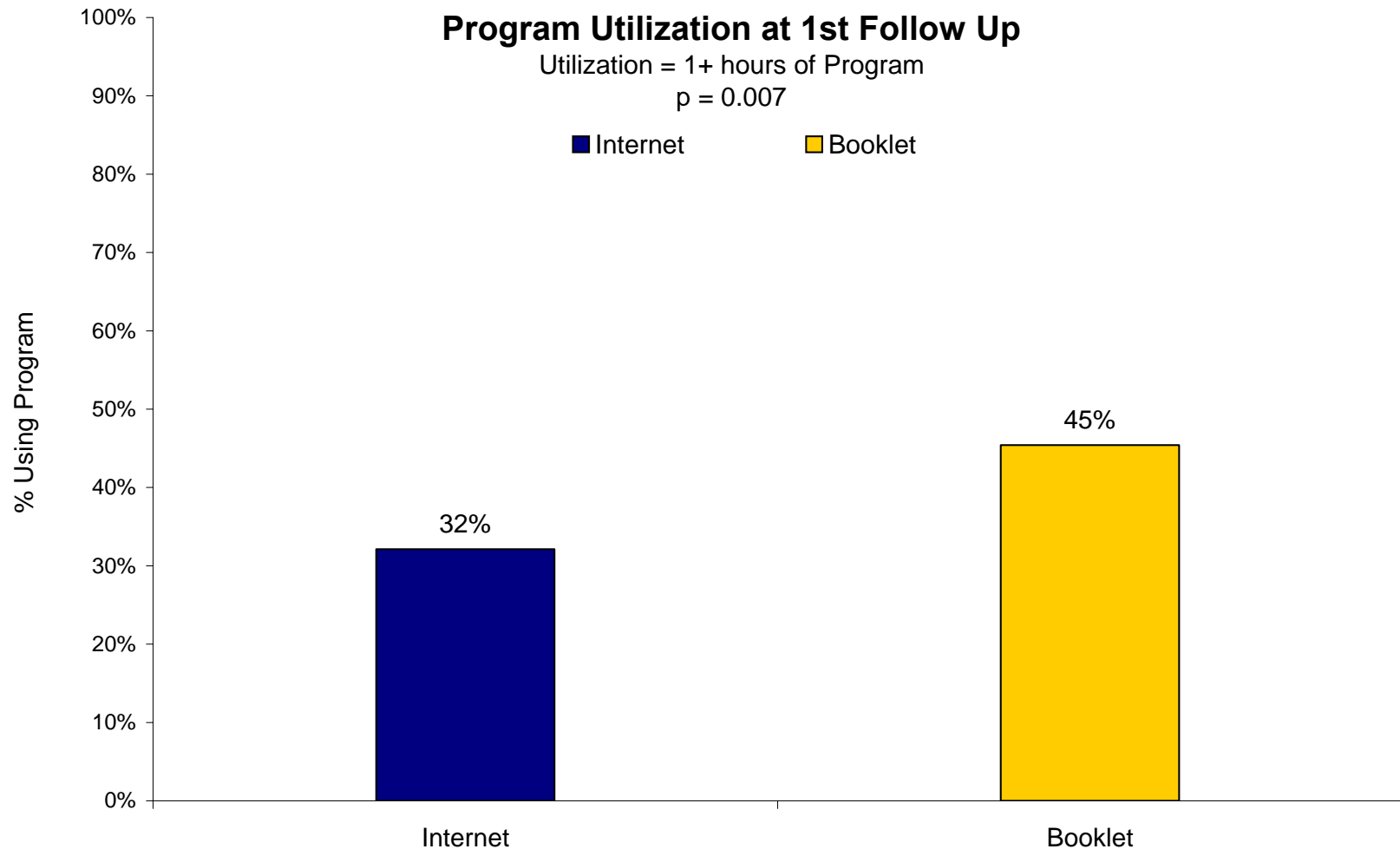
Complication: Program Use

- Not all **used** assigned program
- Examine **Program Use** as potential mediator of treatment effects
 - Self-reported use of assigned program
 - Use: ≥ 1 hour
 - Non-Use: < 1 hour

Program Use



- Use at 1st follow up is better for the booklet



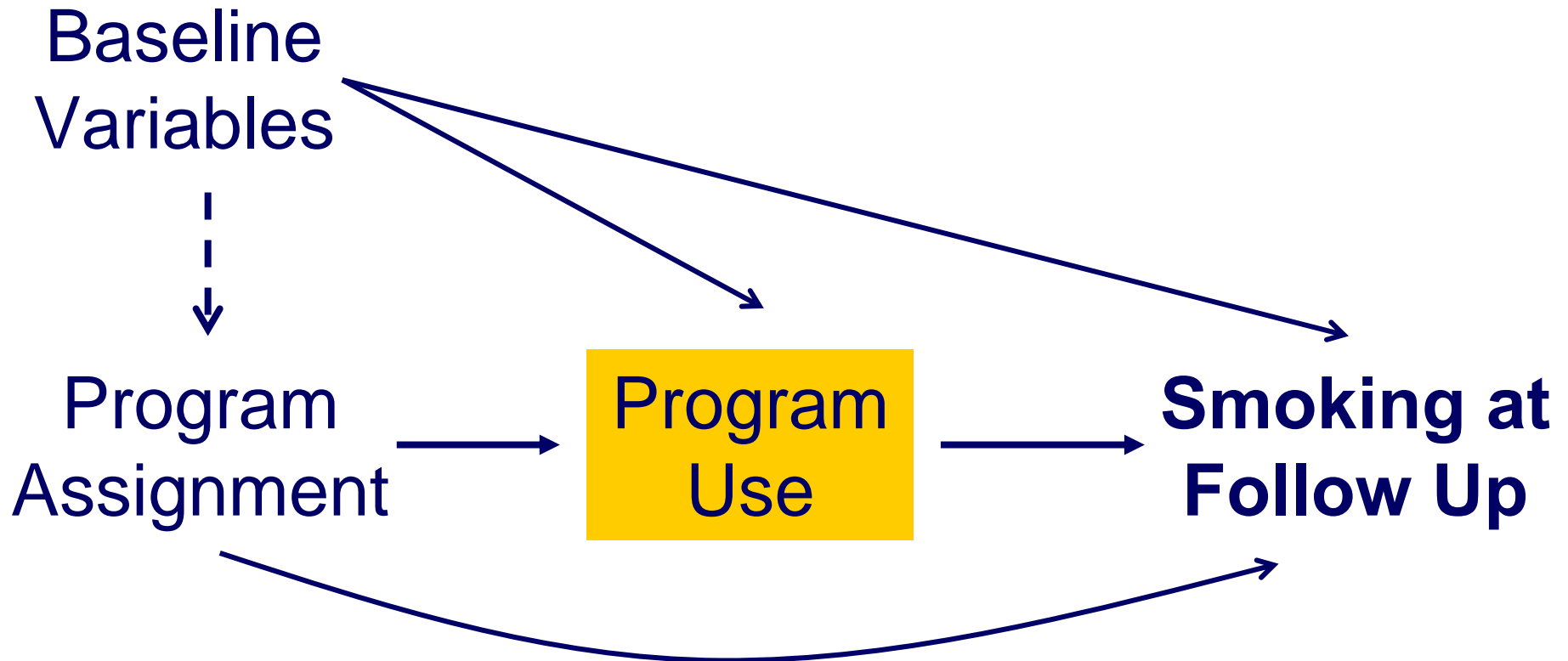
Real Question of Interest



- Is one more likely to succeed at quitting smoking *when using* a self-help internet program than *when using* a self-help booklet?



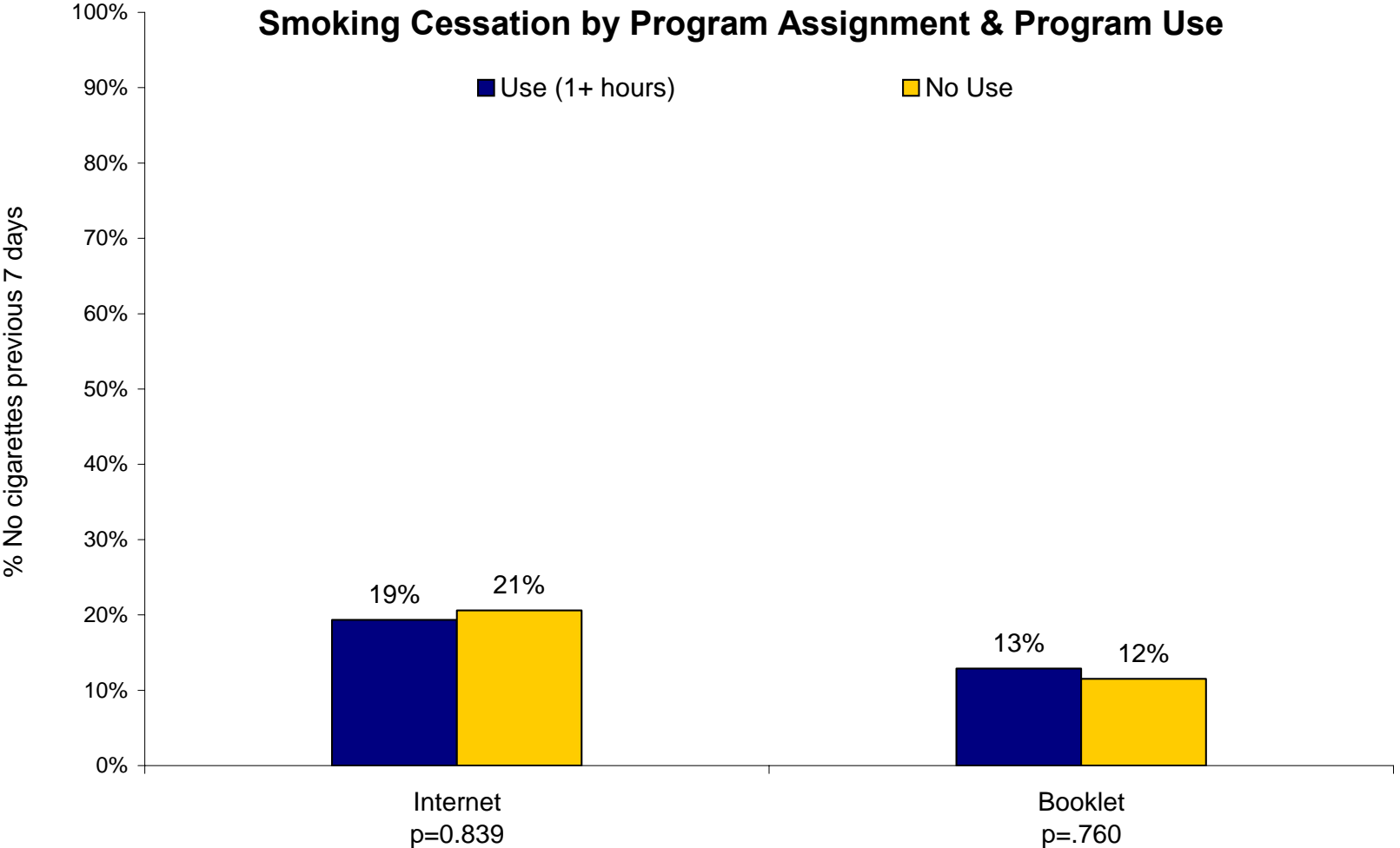
Causal Model





Program Use & Quitting

- In simple comparison, program use does not seem to matter.





Potential Model

program assignment
program use
age
education
language
marriage
intent to quit
wait for payment
addiction

assignment * use
assignment * age
assignment * education
assignment * language
assignment * marriage
assignment * intent to quit
assignment * payment
assignment * addiction

intent * addiction
addiction * intent
payment * intent
payment * addiction
intent * language

age * education
age * language
education * marriage
education * language

use * age
use * education
use * language
use * marriage
use * intent to quit
use * payment
use * addiction

MSM Step 1: Model Treatment



1. Model the **treatment (program use)** using all covariates hypothesized to affect outcome (smoking)
 - Probability someone reported program use, given their observed covariates
 - Each subject is then weighted by the **inverse** of that probability
 - Reweights the observed population to reflect what **would have** occurred if the subjects were randomized



Treatment Model

$$\begin{aligned} \text{Logit}(\mathbf{Program\ Use}) = & -2.0 + 0.4 (\text{program}) - 0.03 (\text{age}) + \\ & 1.9 (\text{education}) + 3.7 (\text{language}) - 1.2 (\text{marriage}) + 1.5 \\ & (\text{payment}) + 0.6 (\text{addiction}) + 0.001 (\text{program*age}) + \\ & 0.3 (\text{program*education}) + 0.1 (\text{program*language}) - \\ & 0.9 (\text{program*marriage}) + 1.4 (\text{program*payment}) + \\ & 1.2 (\text{program*addiction}) - 0.3 (\text{age*education}) - 1.4 \\ & (\text{age*language}) + 1.0 (\text{age*marriage}) - 1.0 \\ & (\text{age*payment}) - 0.8 (\text{age*addiction}) - 1.4 \\ & (\text{education*language}) - 0.4 (\text{education*marriage}) - 1.1 \\ & (\text{education*payment}) - 1.2 (\text{education*addiction}) + 0.5 \\ & (\text{language*marriage}) - 2.6 (\text{language*payment}) + 3.8 \\ & (\text{language*addiction}) + 0.6 (\text{marriage*payment}) + 0.7 \\ & (\text{marriage*addiction}) + 0.3 (\text{payment*addiction}) \end{aligned}$$

MSM Steps 2 & 3: Weight Data



2. Calculate a **weight** for each person

- Plug in the values of all variables
- Sum
- Exponentiate (because logistic regression)
- Take inverse of (1 over) that value → “inverse probability weight”

3. Do **weighted regression**

- Using above weights
- Only main variable(s) of interest in the model

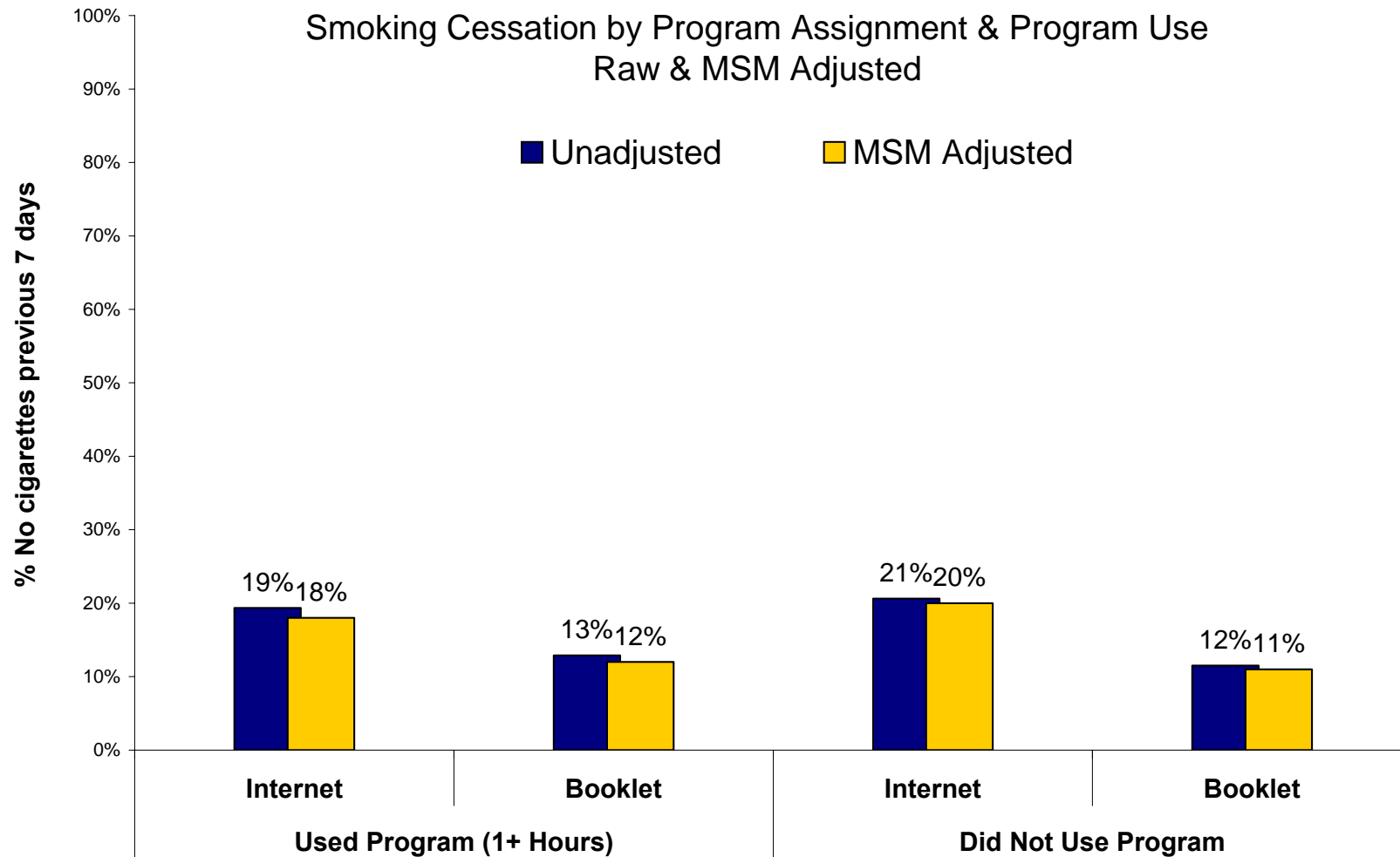


MSM Model

- Outcome = Quit Smoking (no cigarettes previous 7 days)

Variable	Odds Ratio	P value
Intercept	0.24	<.001
Program Assignment <i>Booklet (0) vs. Internet (1)</i>	1.79	0.22
Program Use <i>Not use (0) vs. Use (1)</i>	0.90	0.85
Program Assignment * Use	0.89	0.90

Program Use & Quitting: MSM Adjusted





Limitations of MSMs

- Not always easy to understand
 - Most instructional materials are hard to follow
 - Difficulty accepting “black box” method
- Computing complexities limit application
- Does not mitigate conventional regression assumption of “no unmeasured confounders”
- Marginal effects might mask subgroup differences
 - **Look at your data first**



Summary

- MSMs approximate RCTs with observational data
- The conventional approach is OK when there is
 - Perfect randomization
 - No informative missing data
 - No covariates in the causal pathway
- MSMs provides
 - Direct *causal* estimates, not associations
 - Population (marginal) effect estimates, not independent (conditional) effects
 - Valid estimates in the presence of complex covariate relationships
- MSMs easily handle censoring & multiple treatment mechanisms
 - The impact of different types of censoring (e.g., missing data & no program use) can be addressed in single analysis



Additional Resources

Introductory

- Greenland S, Pearl J, Robins JM. Causal Diagrams for Epidemiologic Research. *Epidemiology* 1999;10(1):37-48.
- Robins JM. Association, causation, and marginal structural models. *Synthese* 1999;121(1-2):151-179.
- Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550-560.
- Maldonado G, Greenland S. Estimating causal effects. *Int J Epidemiol* 2002;31(2):422-9.
- Bodnar LM, Davidian M, et al. Marginal structural models for analyzing causal effects of time-dependent treatments: An application in perinatal epidemiology. *Am J Epidemiol* 2004;159(10):926-34.

More Complex

- Holland, P. Statistics and Causal Inference. *JASA* 1986;81:945-970.
- Pearl J. *Causality*. Cambridge University Press UK, 2001.
- Cole SR, Hernan MA, et al. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. *Am J Epidemiol* 2003;158(7):687-94.
- Joffe MM, Ten Have TR, Feldman HI, Kimmel SE. Model Selection, Confounder Control, and Marginal Structural Models: Review and New Applications. *The American Statistician* 2004;58:272-279.
- Mortimer KM, Neugebauer R, et al. An application of model-fitting procedures for marginal structural models. *Am J Epidemiol* 2005;162(4):382-8.
- Petersen ML, Deeks SG, et al. History-adjusted Marginal Structural Models for Estimating Time-varying Effect Modification. *Am J Epidemiol* 2007;166(9):985-993.

Software

- <http://www.hsph.harvard.edu/causal/software.htm>