The Immediate Post-Launch Phase and Confounding by Indication or Channeling Bias:

How does it affect our ability to conduct CER?

Cynthia J Girman
Executive Director and Head of
Data Analytics & Observational Methods,
Center for Observational & Realworld Evidence
Merck & Co., Inc

Personal or financial relationships in the past 12 months:

- Employed by commercial entity – Merck & Co., Inc.
- Stock ownership or options - Merck & Co., Inc
- Advisory Board membership- Center for Pharmacoepidemiology, University of North Carolina
- Other – Adjunct Associate Professor UNC Epidemiology

Reminder of Definitions

- ***Confounding by indication**** can occur when the underlying diagnosis or other clinical features that trigger use of a certain product also are related to patient outcome
- ***Channeling bias**+* is a type of selection bias, which occurs when a product is claimed to be safe and therefore is used in high risk patients who did not tolerate other products for that indication
- May apply to any non-randomized comparative studies
  - Pharmacoepidemiology safety studies of unintended effects
  - Comparative effectiveness research of intended effects of products or therapeutic regimens

*Brian Strom (editor). Pharmacoepidemiology, 4th ed. Wiley & Sons LTD: West Sussex, England.
+Many researchers extend this definition of channeling bias to generally refer to situations where a drug, device or therapeutic regimen may be more likely used in a patient population with specific characteristics

Examples of Confounding by Indication or Channeling Bias Affecting CER

Propensity scores reflect the conditional probability of the treatment under study given observed characteristics at the time of treatment initiation

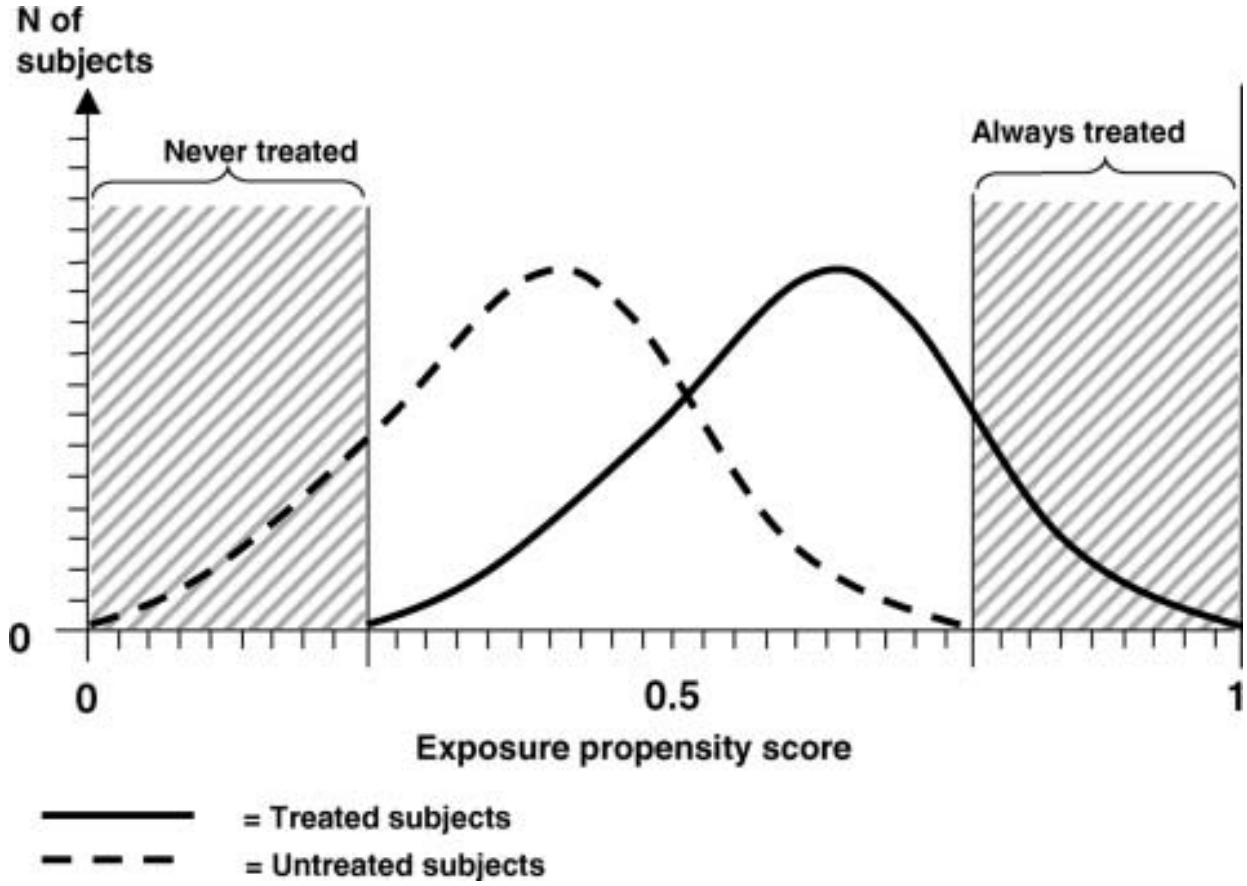
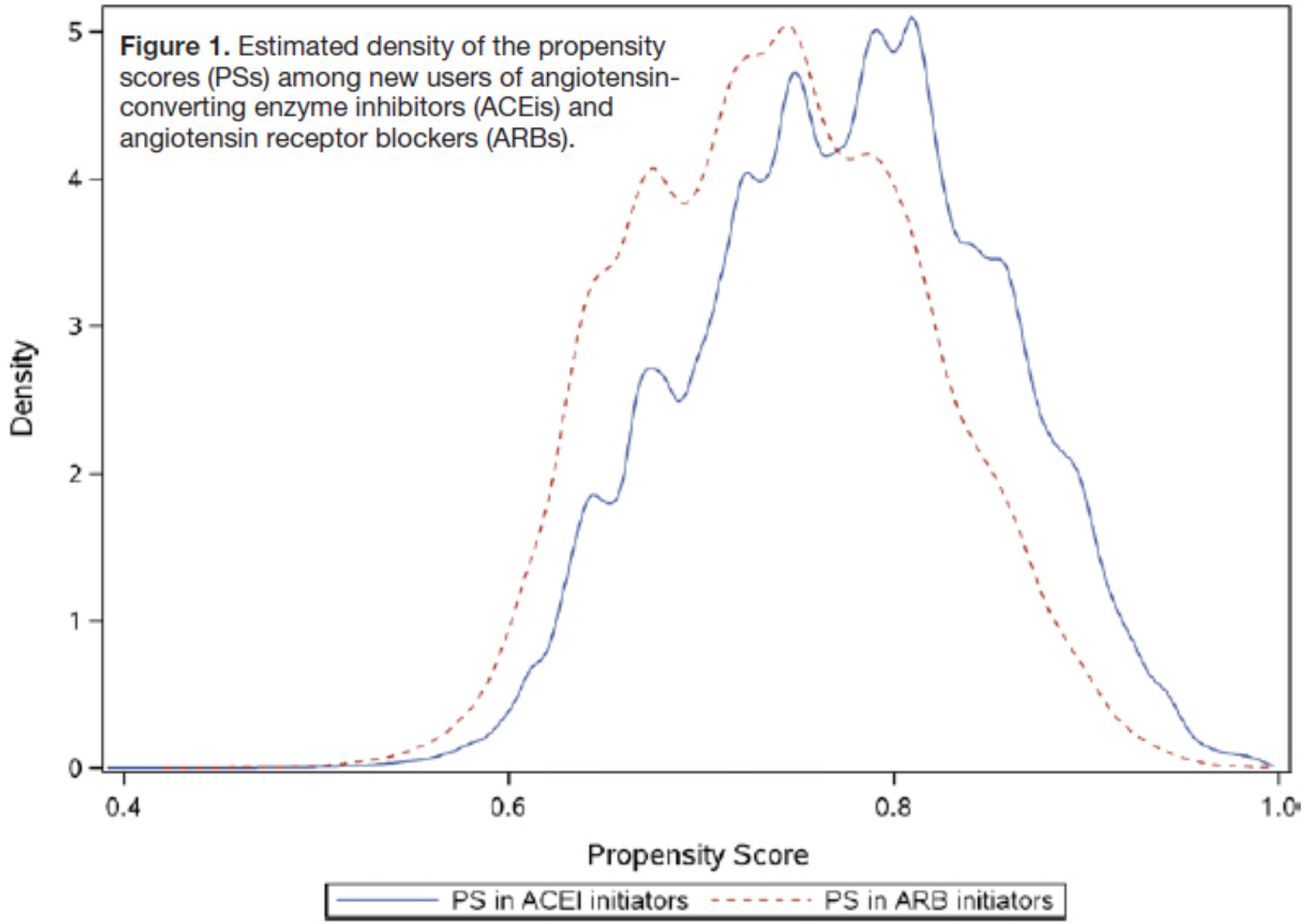


Fig. 2. The non-overlap of the exposure propensity score distribution among treated and untreated study subjects. In this example subjects with very low propensity score are never treated while subjects with very high propensity score are all treated.

From Glynn RJ, Schneeweiss S, Stürmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. *Basic & Clinical Pharmacology & Toxicology* 2006; 98: 253-259.

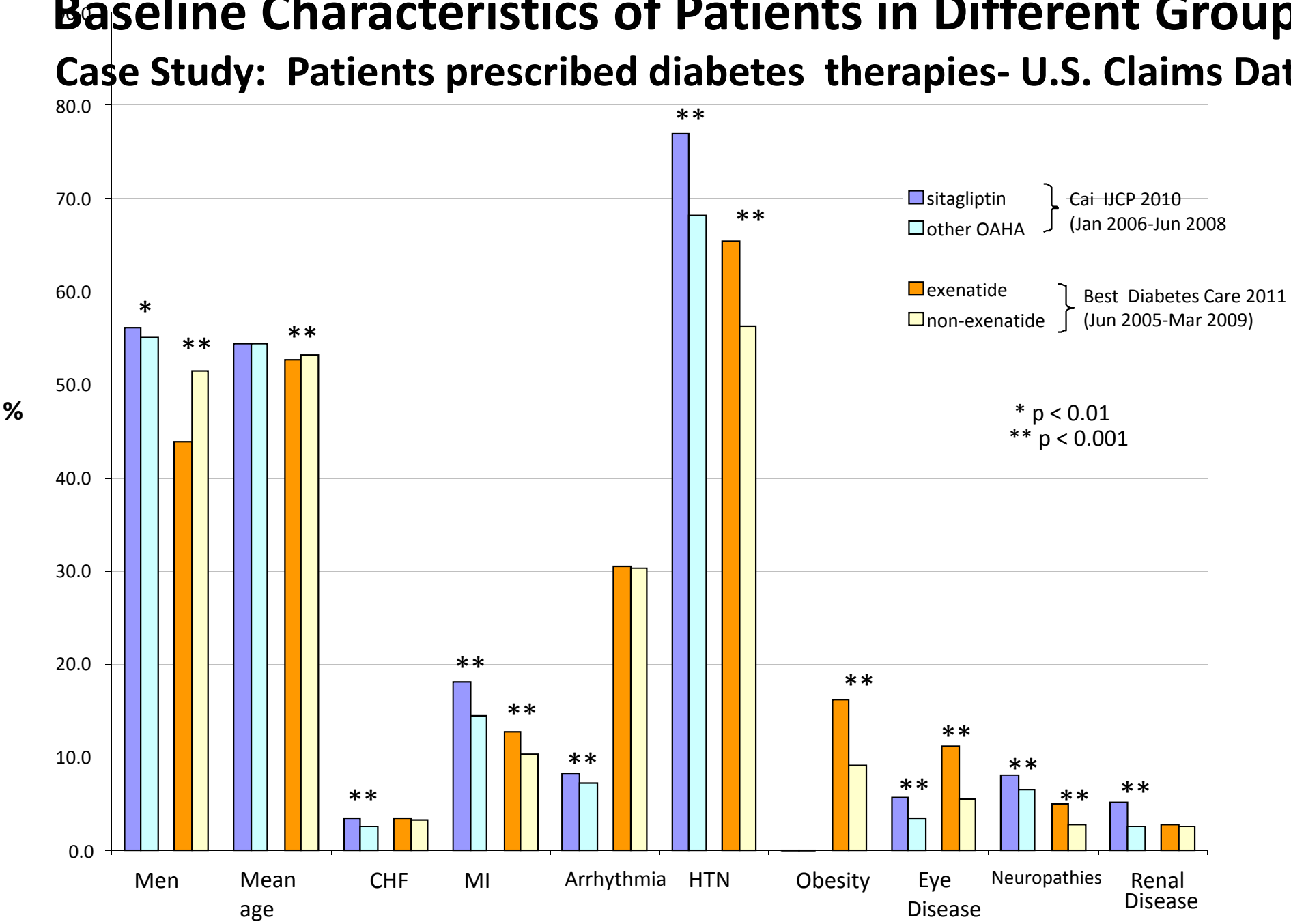
Example of Substantial Overlap in Propensity Scores Between Treatment Groups



From Brookhart MA, Wyss R, Layton JB, Stürmer T. Propensity score methods for confounding control in nonexperimental research. *Circulation: Cardiovascular Quality and Outcomes*. 2013; 6: 604-611.

Baseline Characteristics of Patients in Different Groups

Case Study: Patients prescribed diabetes therapies- U.S. Claims Data



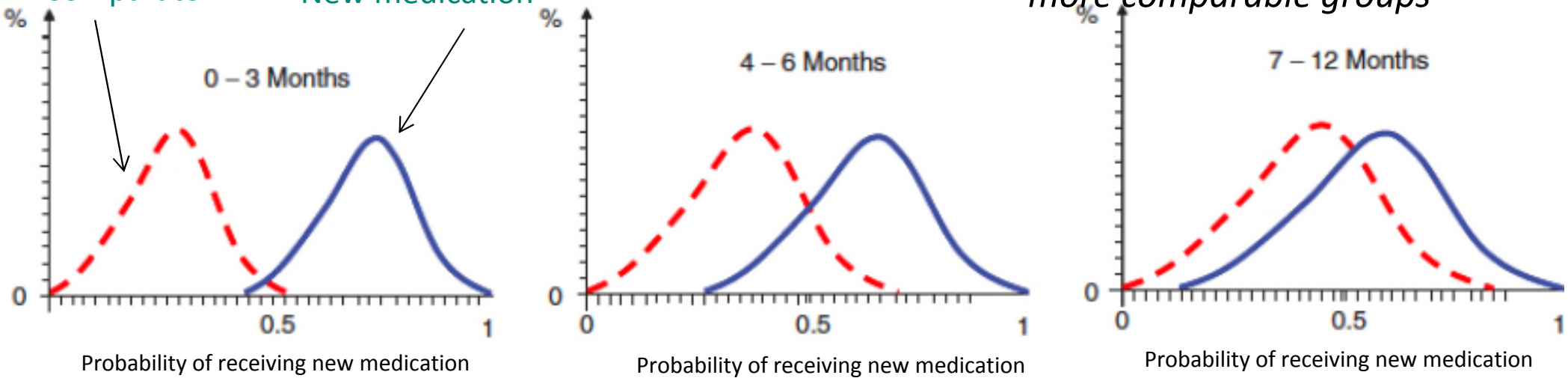
Cai B, Katz L, Alexander CM, Williams-Herman D, Girman CJ. *Int J Clin Pract*. 2010 Nov;64(12):1601-8.

Best JH, Hoogwerf B, Herman WH, Pelletier EM, Smith DB, Wenten M, Hussein MA. et al. *Diabetes Care* 2011; 34: 90-95.

Channeling of Patients to Certain Therapies Often Changes Over Time with Newly Marketed Medications

- When a new medication comes onto the market, sicker patients may receive the new medication at first – over time, patients initiating the new medication may look more like patients initiating a comparator

Increased overlap indicates more comparable groups



→ Increasing time since initial marketing of new medication →

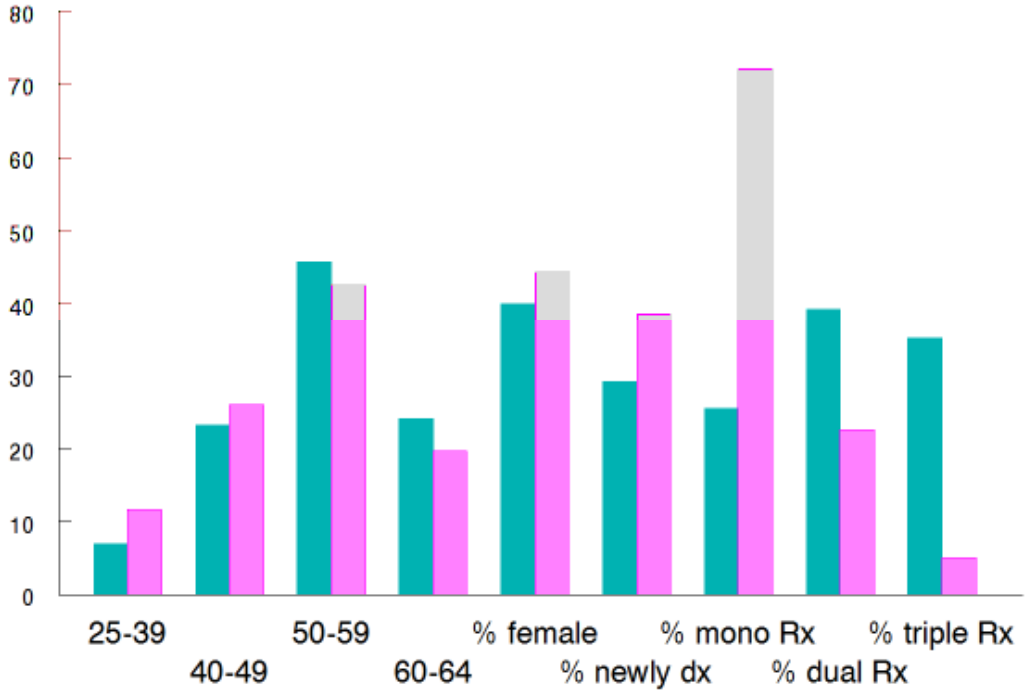
Ref: Schneeweiss et al. *Clin Pharmacol & Ther* 2011; 90: 777-790

Differences in Baseline Characteristics May Persist for Years After Launch: Sitagliptin vs Other OHA

2006-2007

Sitagliptin N=11,550

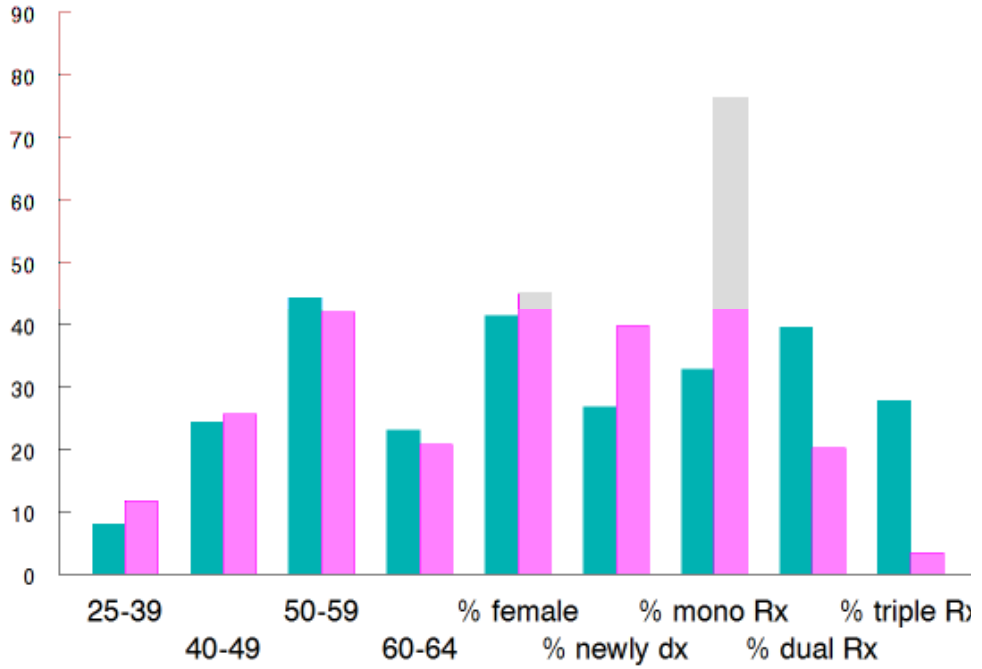
Other OHAs N=106,122



2008-2010

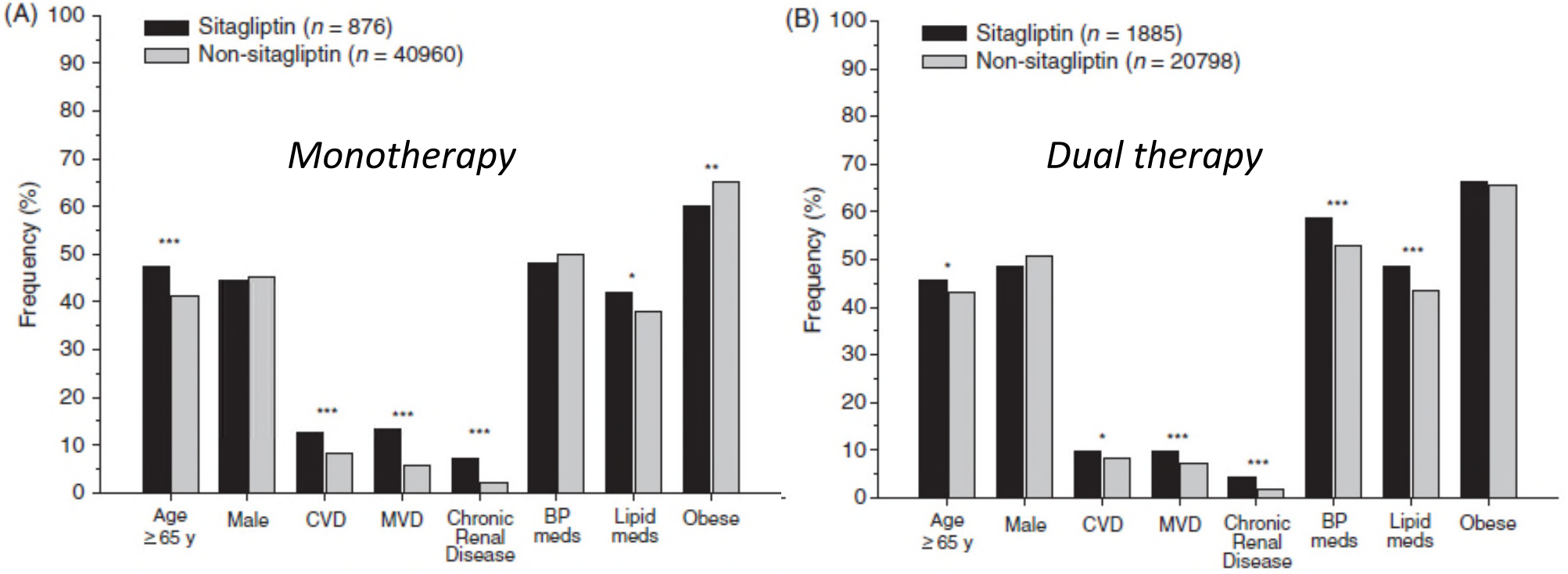
Sitagliptin N=31,085

Other OHAs N=129,117



Brodovicz KG, Kou TD, Alexander CM, O'Neill EA, Senderak M, Engel S, Girman CJ. Int J Clin Prac 2013; 67: 449-454.

Baseline Characteristics of Patients Prescribed Sitagliptin and Other Antihyperglycemic Agents– GE Centricity

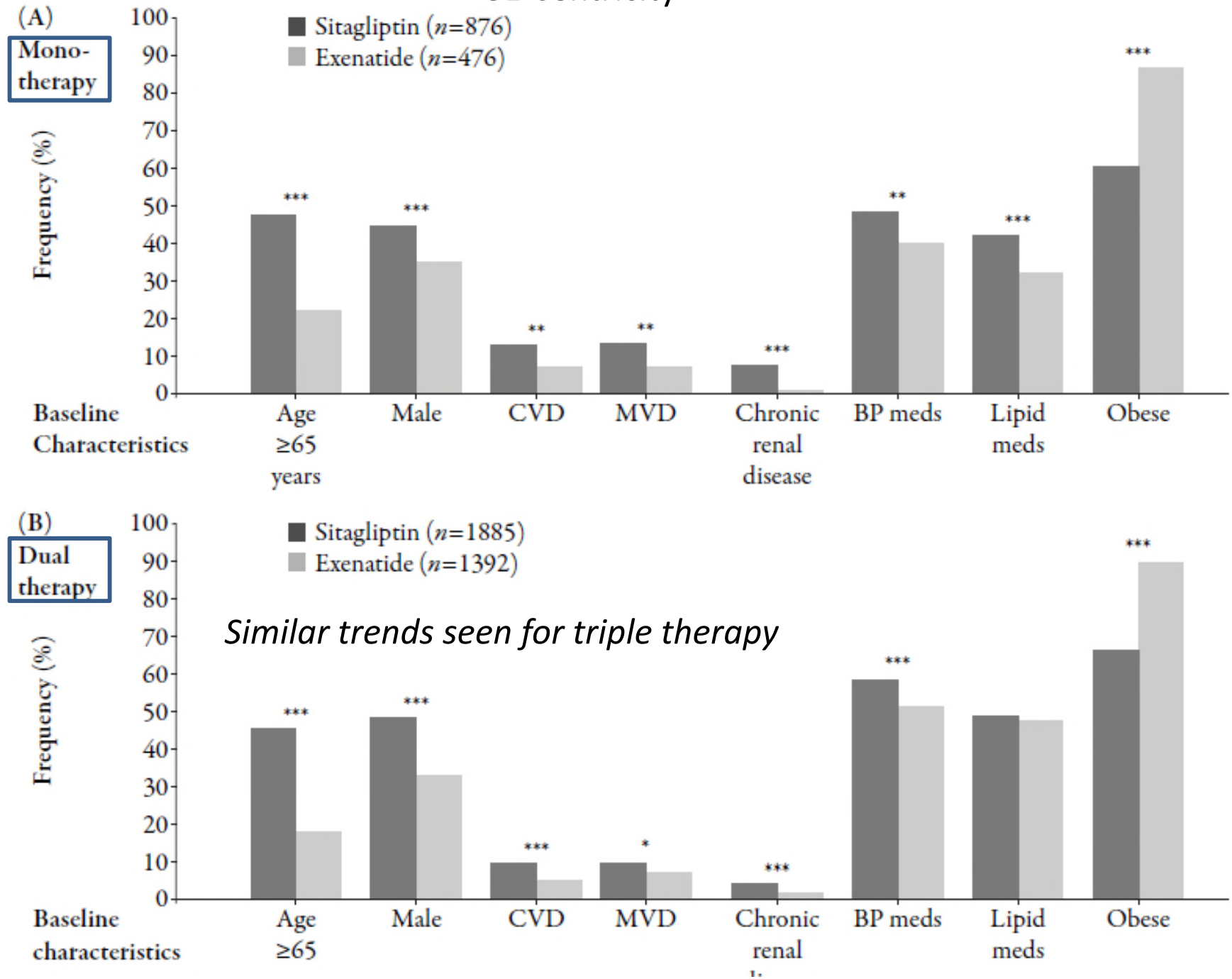


CVD-cardiovascular disease-related conditions; MVD=microvascular disease-related conditions, BP-blood pressure; meds=medications

Zhang Q, Mavros P, Engel SS, Davies MJ, Yin D, Radican L, Rajagopalan S. *Curr Med Res Opin* 2010; 26:1697–703.

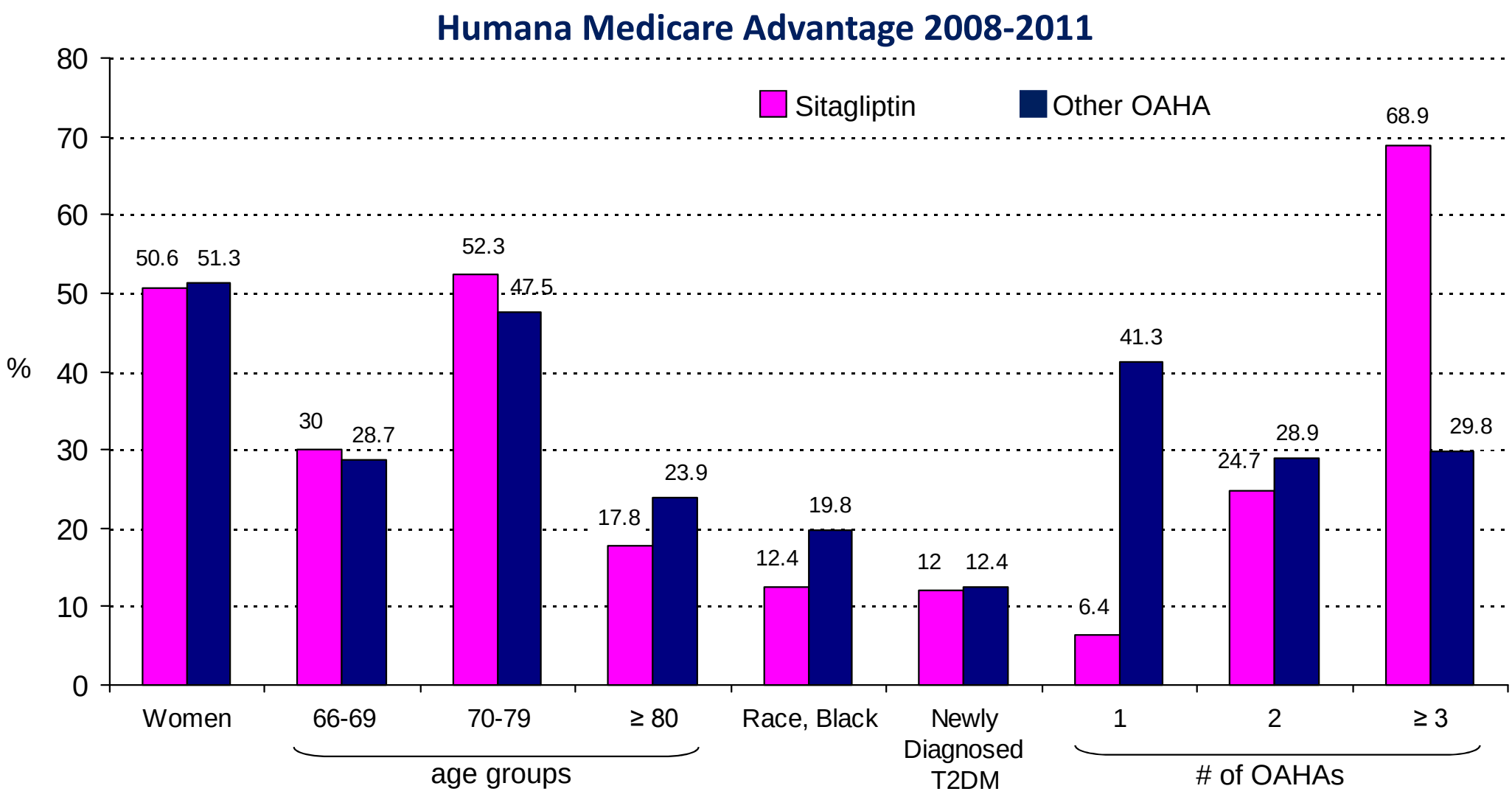
Baseline Differences Between Patients Prescribed Sitagliptin and Exenatide 2006-2008

GE Centricity



Zhang Q, Rajagopalan S, Mavros P, Engel SS, Davies MJ, Yin D, Radican L. Adv Ther 2010; 27: 223-232.

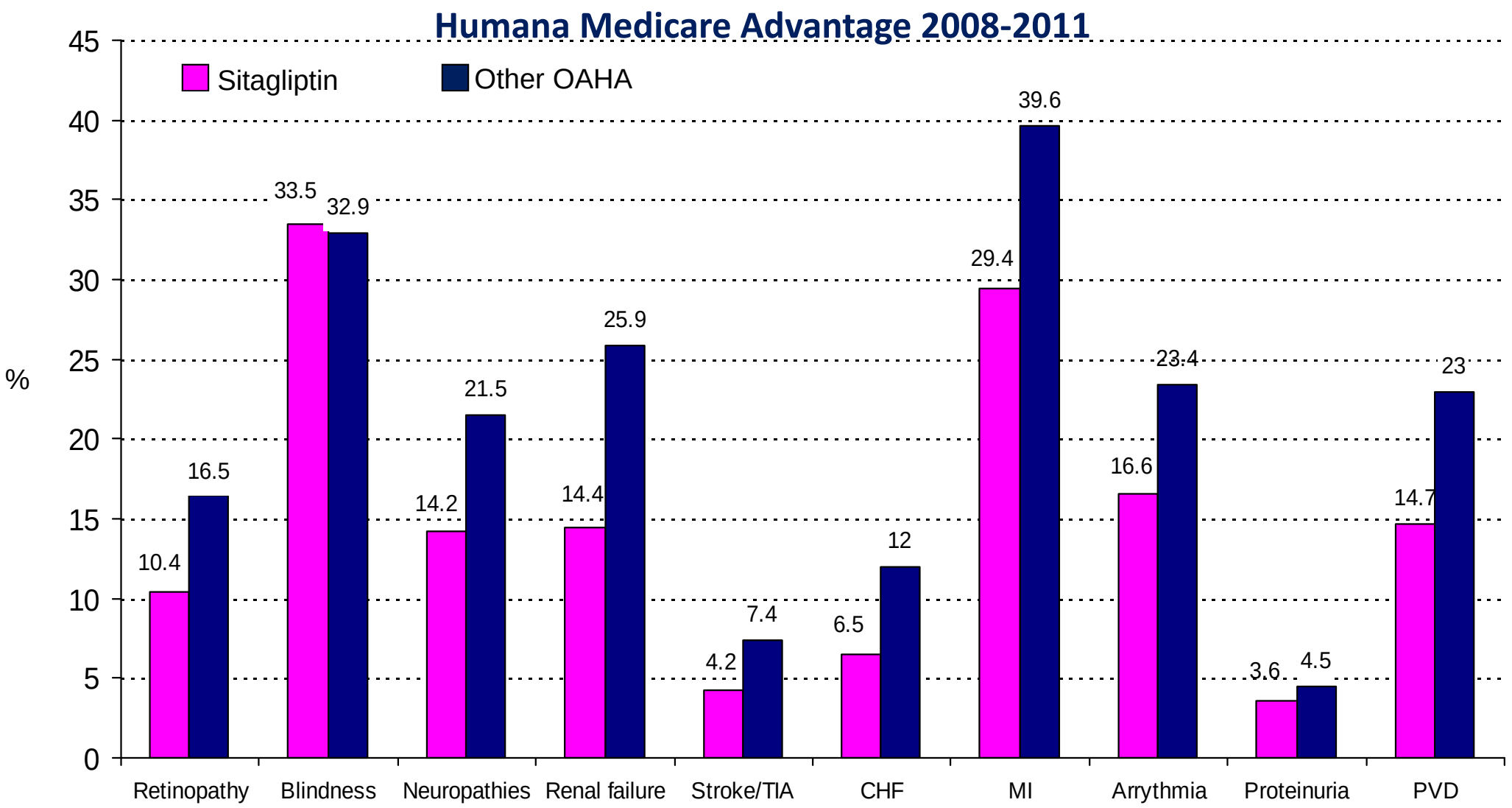
Baseline Characteristics of Elderly patients (≥65 years) initiating Sitagliptin or Other Oral Antihyperglycemic Agents (OAHAs)



Adjusted OR (95% CI)	1.02 (0.97, 1.07)	Reference	**1.14 (1.08, 1.21)**	0.98 (0.91, 1.05)	**0.68 (0.64, 0.73)**	0.94 (0.87, 1.02)	Reference	**5.15 (4.68, 5.66)**	**13.21 (12.07, 14.44)**

Brodovicz KG, Kou TD, Engel SS, Alexander CM, O'Neill EA, Girman CJ. Diabetes 2012; 61(S1): A350 (ADA)
Brodovicz KG et al, PDS 2012; 21(S3): 69.

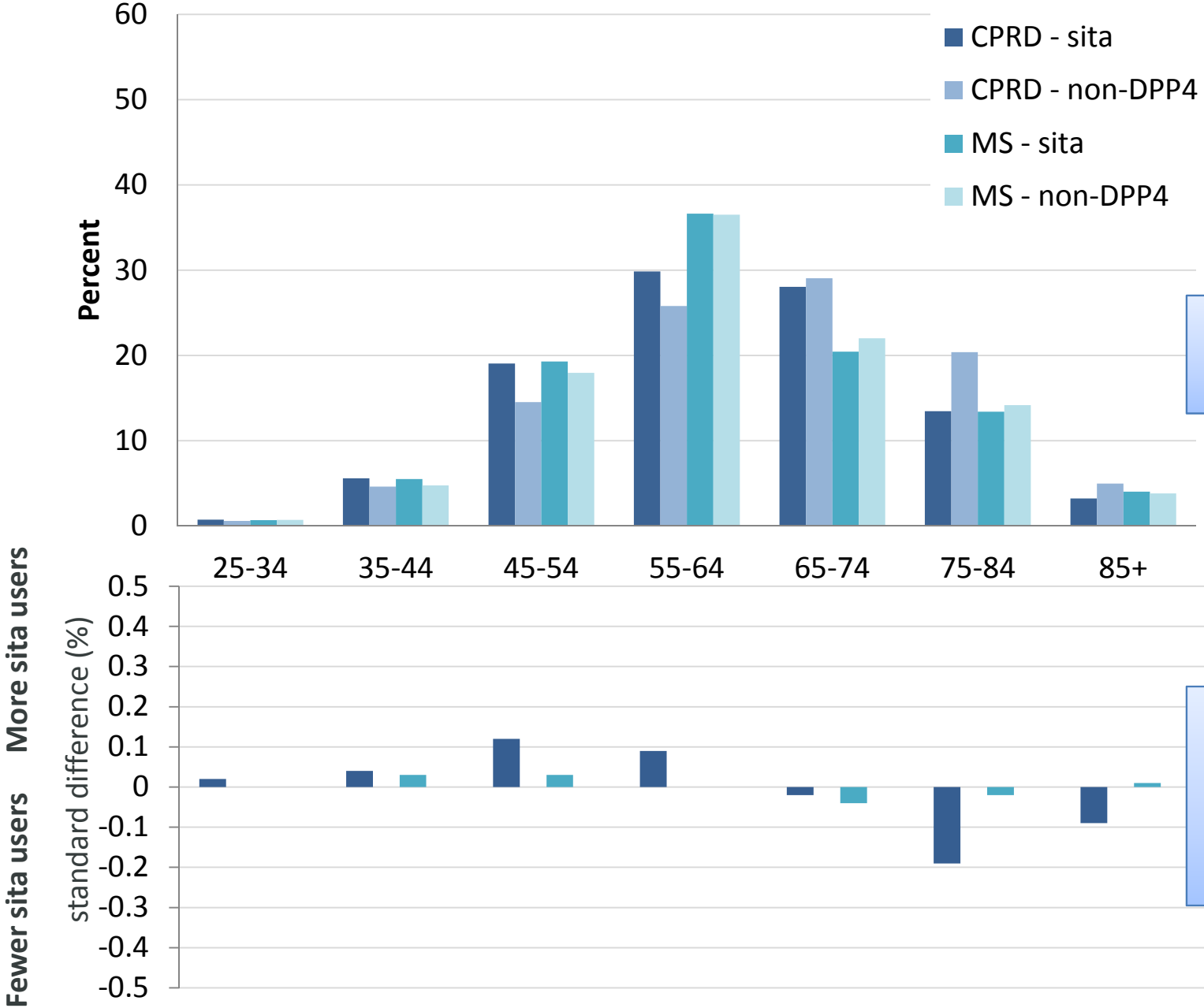
Diabetes-Related Complications and Comorbidities Prior to Initiation of Sitagliptin or Other Oral Antihyperglycemic Agents in Elderly Patients



Adjusted OR (95% CI)	**0.71 (0.65, 0.77)**	**1.15 (1.09, 1.21)**	**0.72 (0.67, 0.77)**	**0.85 (0.79, 0.91)**	**0.78 (0.70, 0.88)**	**0.89 (0.80, 0.98)**	**0.85 (0.80, 0.90)**	0.96 (0.89, 1.03)	1.07 (0.94, 1.22)	**0.82 (0.76, 0.88)**

Brodovicz KG, Kou TD, Engel SS, Alexander CM, O'Neill EA, Girman CJ. Diabetes 2012; 61(S1): A350 (ADA)
Brodovicz KG et al, PDS 2012; 21(S3): 69.

Dual Combination therapy: Age distribution and standardized differences at initiation of sitagliptin or other non-DPP-4i oral antihyperglycemic agents

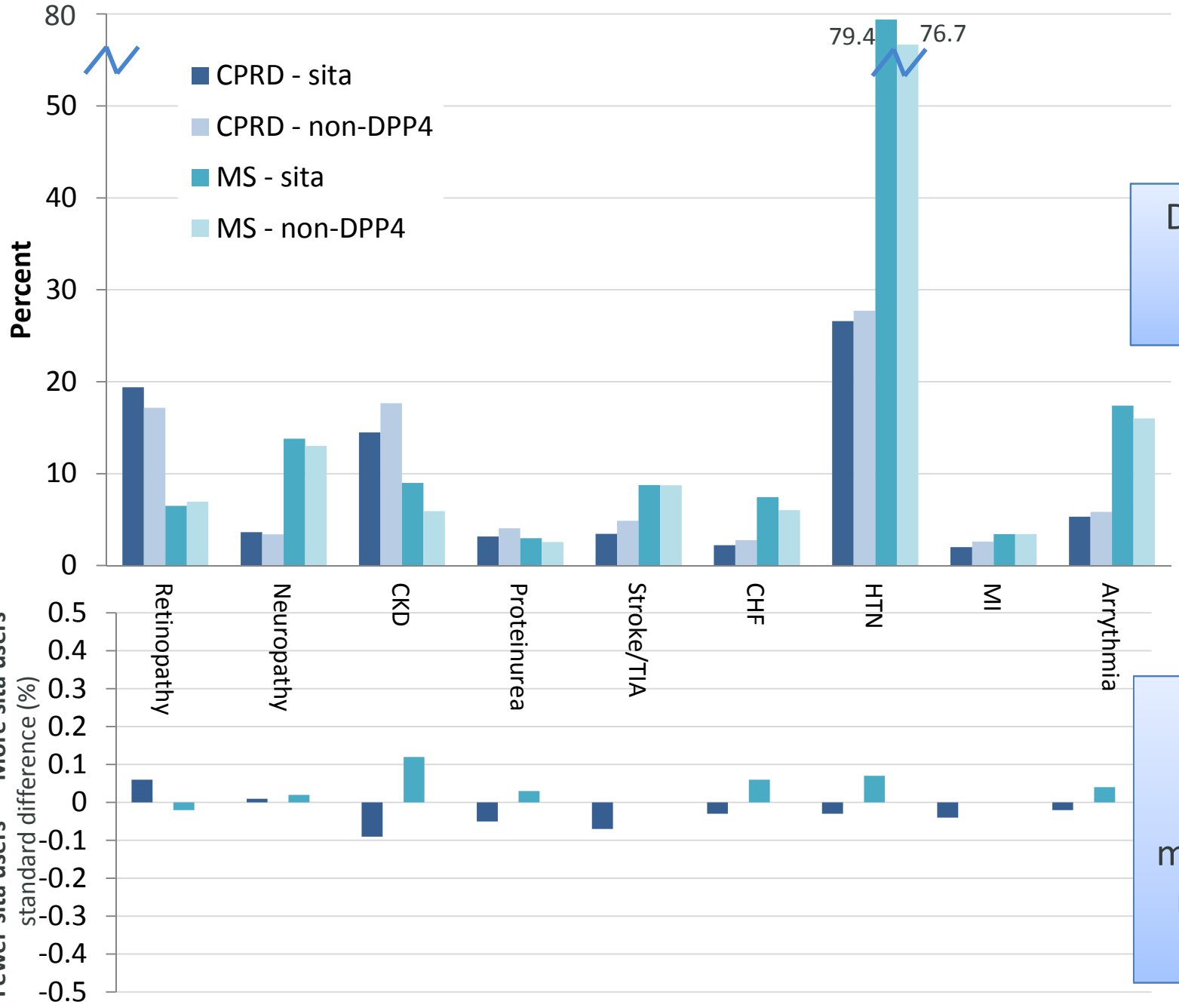


Similar age distributions across databases

More sitagliptin use among younger patients and less among older patients

Ref:

Dual Combination therapy: Diabetes-related complications and comorbidities in the 5 years before including sitagliptin or other non-DPP-4i oral antihyperglycemic agents (distribution and standardized differences)

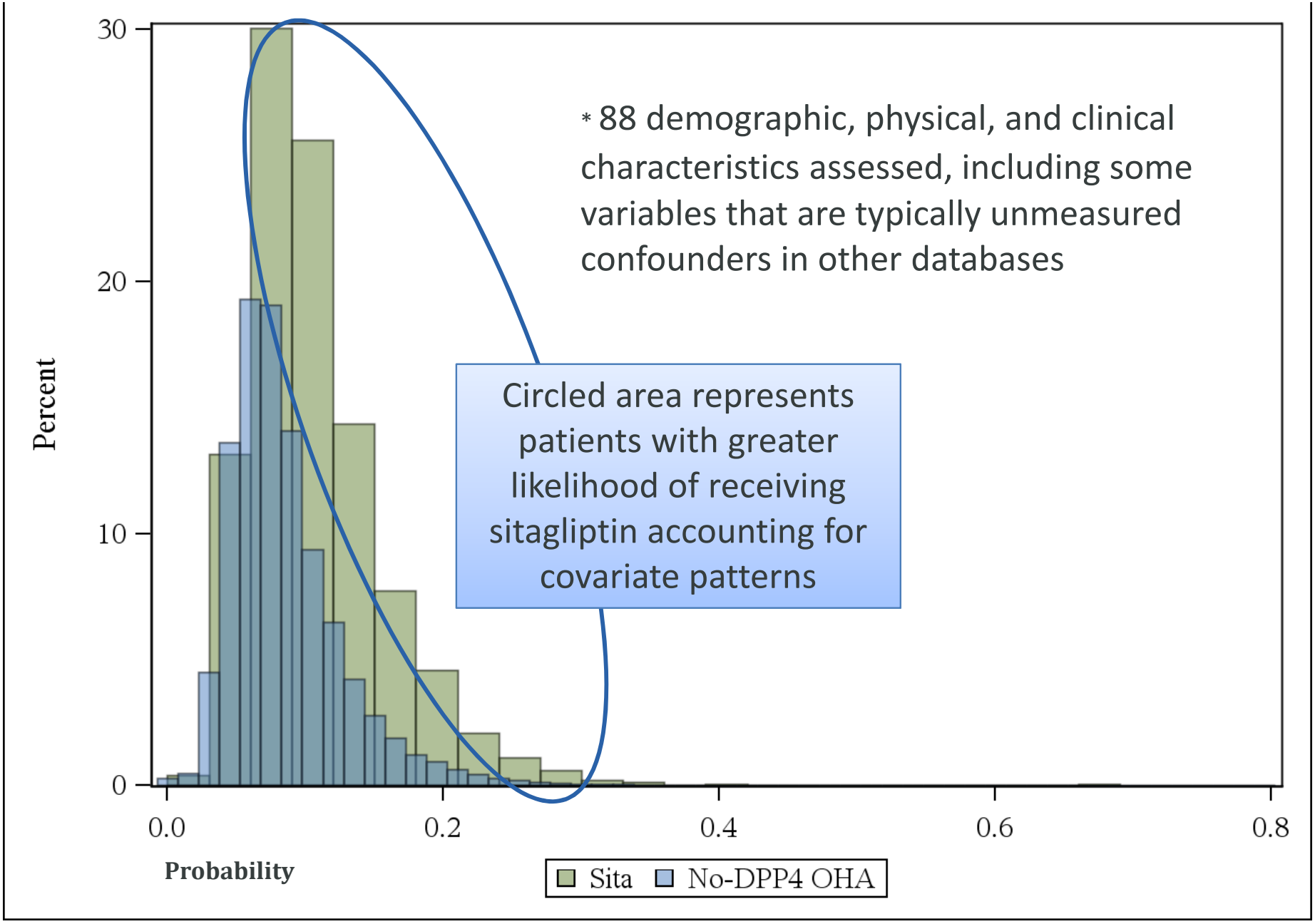


Differences in capture of variables across databases

Differences between patients starting sitagliptin vs other medications – direction not consistent across databases

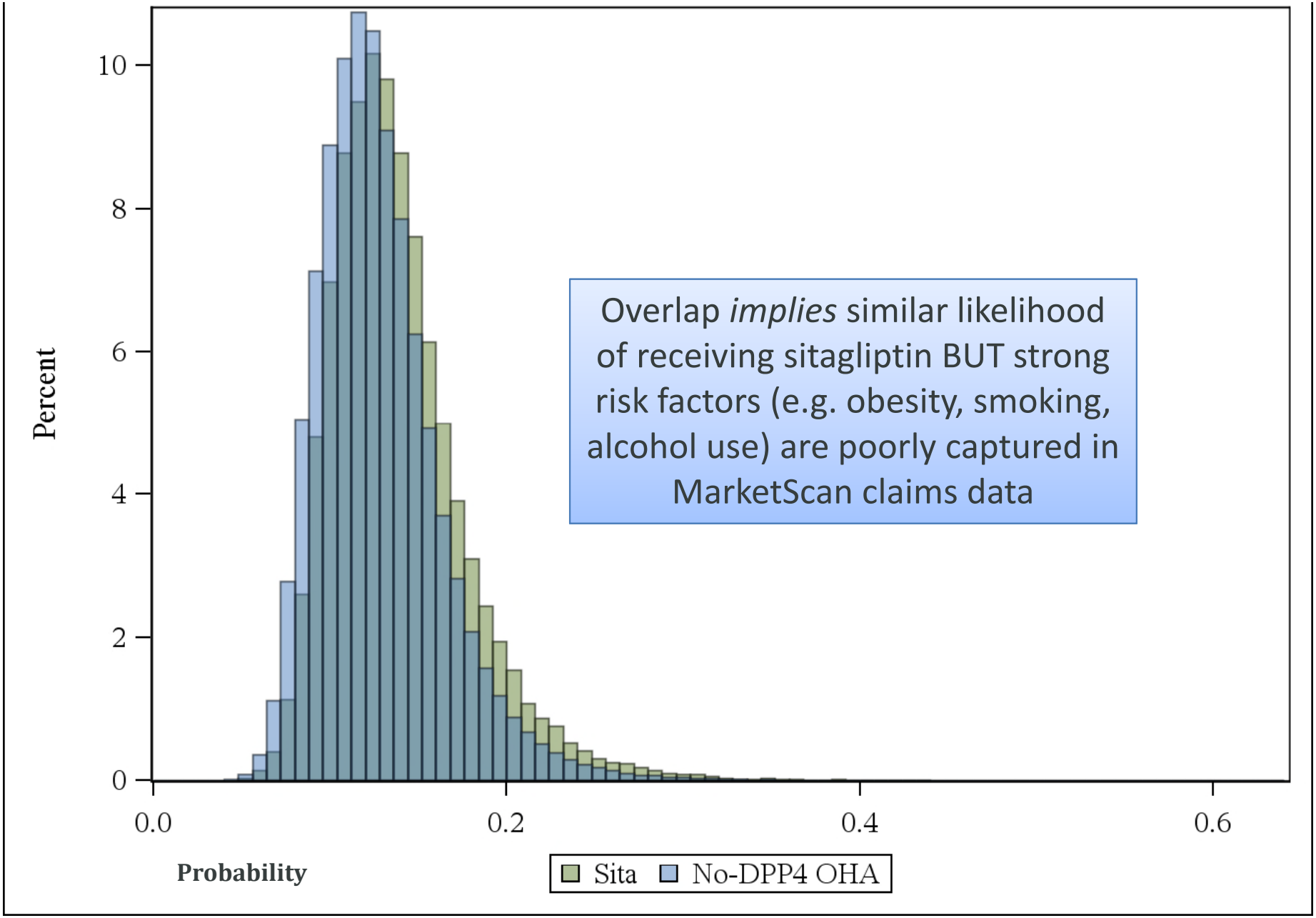
Ref:

Dual Combination Therapy: Probability distribution of sitagliptin vs. non-DPP4i oral use *a priori* propensity score model* - **CPRD**



Ref:

Dual Combination Therapy: Probability distribution of sitagliptin vs. non-DPP4i oral use *a priori* propensity score model* - **MarketScan**



* 106 demographic, physical, and clinical characteristics assessed

Ref:

Summary

- Confounding by indication can be identified for measured confounders by examining baseline characteristics
  - Understanding whether the potential factors are related to both treatment and outcome is critical
  - This is often a subjective assessment
- Confounding by indication or channeling bias may change over time and may vary by subgroup (e.g., elderly)
- Graphing the propensity score or disease risk score overlap can help understand the feasibility of conducting a comparative study with minimal residual confounding
  - However, to understand drivers of the potential baseline differences, summary of individual covariates are useful

Additional Steps for Feasibility and Pre-Identifying Sensitivity Analyses Before Launching Comparative Effectiveness Research

Cynthia J Girman
Executive Director and Head of
Data Analytics & Observational Methods,
Center for Observational & Realworld Evidence
Merck & Co., Inc

Exchangeability Assumption of CER

- Exchangeability[1-2] between treatment groups in terms of covariate distributions
  - Implied is that no clinical or demographic feature distinguishes recipients of one therapy vs the other
  - This can also be described as ‘equipoise’ where strong observed similarity exists between the types of patients receiving two therapies

Greenland & Morgenstern. Confounding in health research. Ann Rev Public Health 2001;22:189-212.
Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. Int J Epidemiol 1986;15:413-9.

Propensity Scores to Balance Groups

- Propensity scores (PS) to balance groups in comparative observational studies increased substantially in last decade
  - Conditional probability of treatment given all the confounders
    - Derived by modeling treatment received (usually with logistic regression) as a function of confounders
  - With a correctly specified PS model and large N, subjects with same PS have similar measured covariate distributions, regardless of treatment group
  - Matching, stratifying, inverse probability weighting of propensity scores are often used to balance the treatment groups for the *measured* covariates
    - This does not hold for unmeasured covariates and confounders

Overlap in Propensity Scores to Determine Feasibility of CER

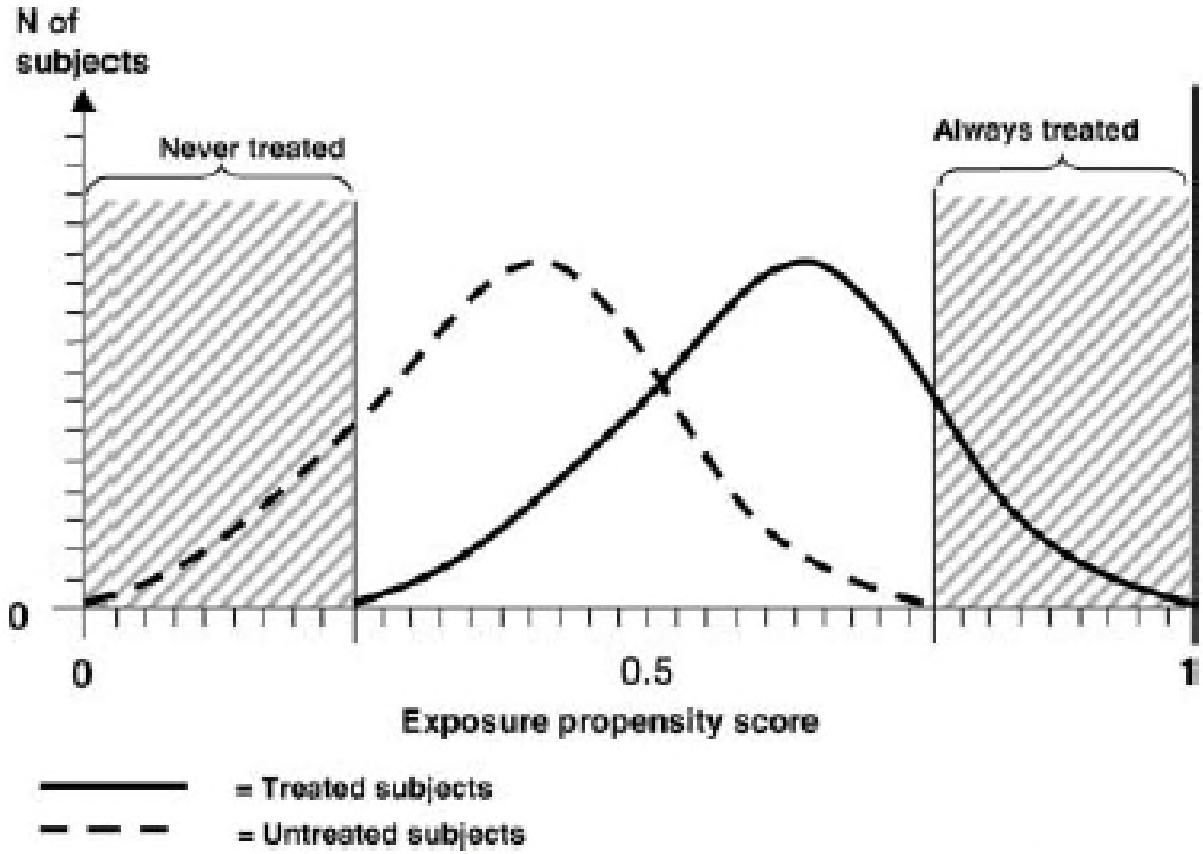
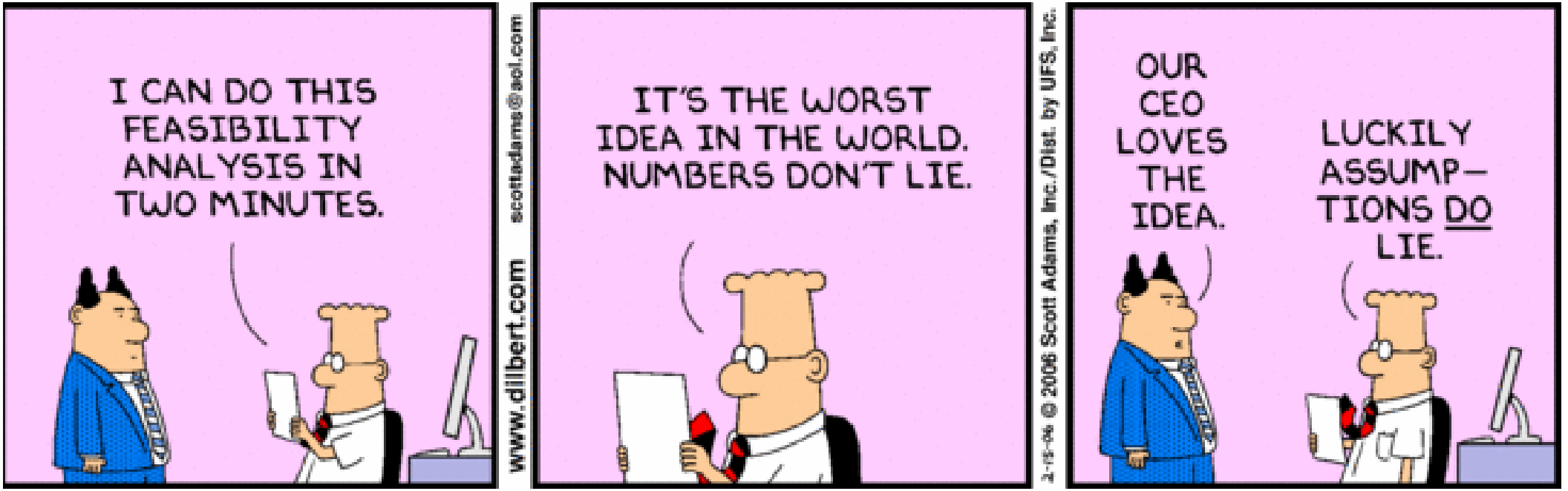


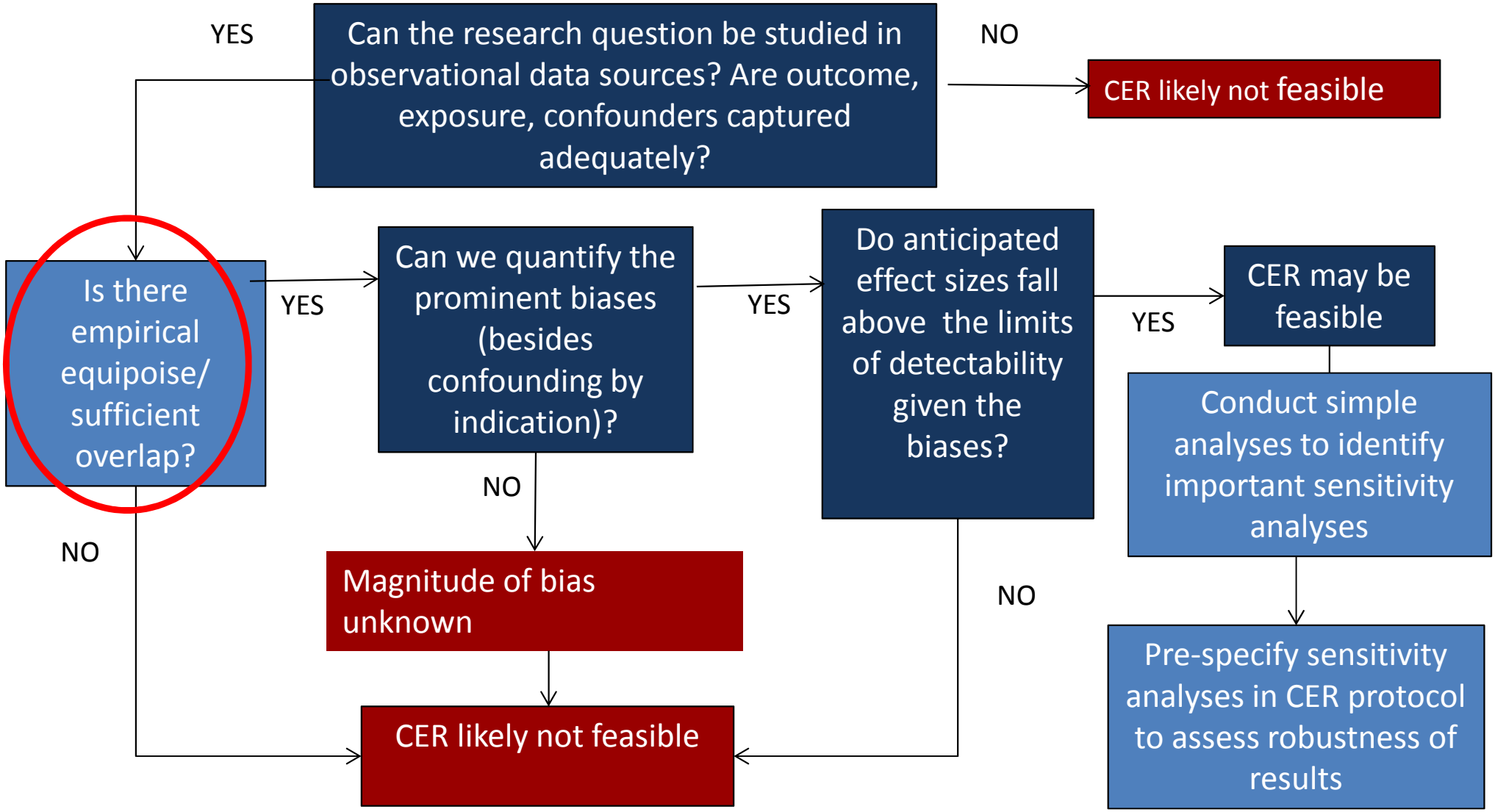
Fig. 2. The non-overlap of the exposure propensity score distribution among treated and untreated study subjects. In this example subjects with very low propensity score are never treated while subjects with very high propensity score are all treated.

Glynn RJ, Schneeweiss S, Stürmer. Indications for Propensity Scores and Review of their Use in Pharmacoepidemiology. *Basic & Clin Pharmacology & Toxicology* 2006; 98: 253-259.

Feasibility Analysis vs Assumptions



Pre-Specify Prior to CER Launch



Girman CJ, Faries D, Ryan P, Rotelli M, Belger M, Binkowitz B, O'Neill R. Pre-Study Feasibility and Identifying Sensitivity Analyses for Protocol Pre-Specification in Comparative Effectiveness Research. *J Comparative Effectiveness Research* 2014; 3: 259-270.

Walker AM *et al*. *Comparative Effectiveness Research* 2013; 3: 11-20

Empirical Equipoise

- **Empirical Equipoise** – A balance of opinion in the prescribing community about what might be the best treatment for a given class of patients
  - Overall, similar patients are prescribed two different therapies
- Equipoise simulates an RCT by making treatment groups "balanced" across measured factors that may predict or be related to the outcome
- It is assumed that there is minimal to no unmeasured confounding

Walker AM *et al*. *Comparative Effectiveness Research* 2013; 3: 11-20.

How to Assess Empirical Equipoise

1. Develop a propensity probability using logistic regression with treatment group (intervention vs comparator) as dependent variable and confounders as independent variables
2. Create preference score ranging 0 - 1 from the predicted values of logistic regression (e.g. probability distribution score)

   $\ln \{F / (1-F)\} = \ln (S/(1-S) - \ln (P/(1-P))$, where

   F=preference score, S=propensity score, P=prevalence of exposure to treatment A
3. Examine preference distributions separately by treatment group
4. Assess whether clinical equipoise is present, if at least half of the distributions lie between 0.3 and 0.7

Walker AM *et al*. *Comparative Effectiveness Research* 2013; 3: 11-20.

Feasibility Assessment

- Use graphical tools to assess overlap in either propensity scores or preference rankings to allow a rationale assessment of equipoise
- Walker et al proposed that preference rankings in from 0.3 – 0.7 reflect sufficient equipoise such that results of CER may be interpretable
  - While arbitrary, the degree of overlap is important and most researchers are quite subjective
- Such tools could be used to assess feasibility of CER and prevent the launch of CER that may be hopelessly confounded and uninterpretable

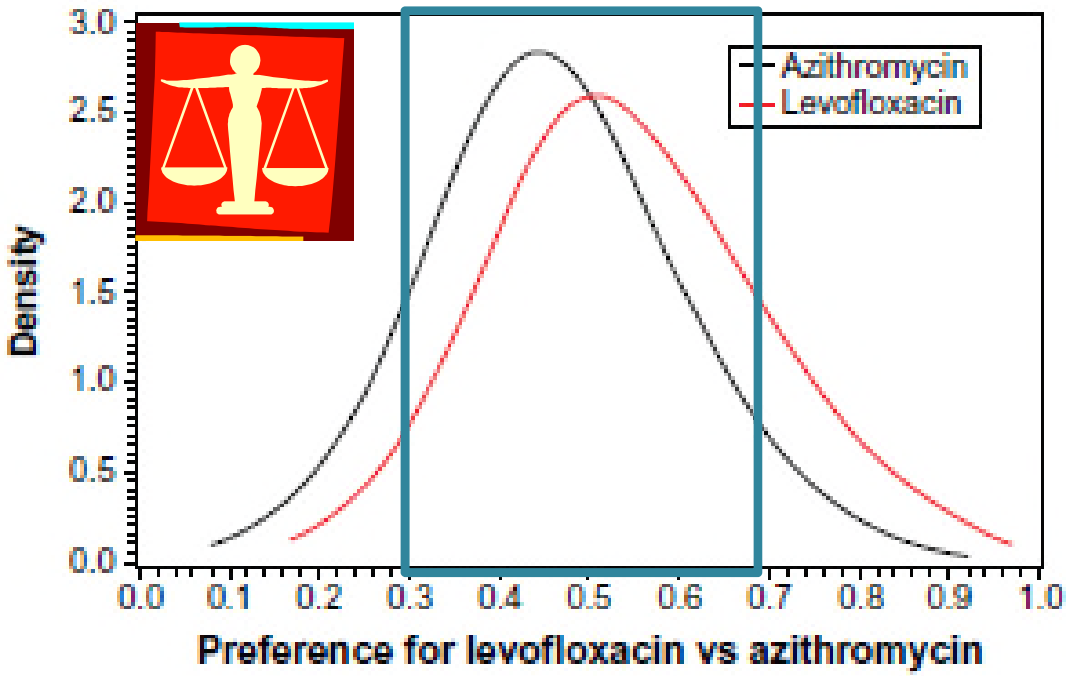
Walker AM *et al*. *Comparative Effectiveness Research* 2013; 3: 11-20.

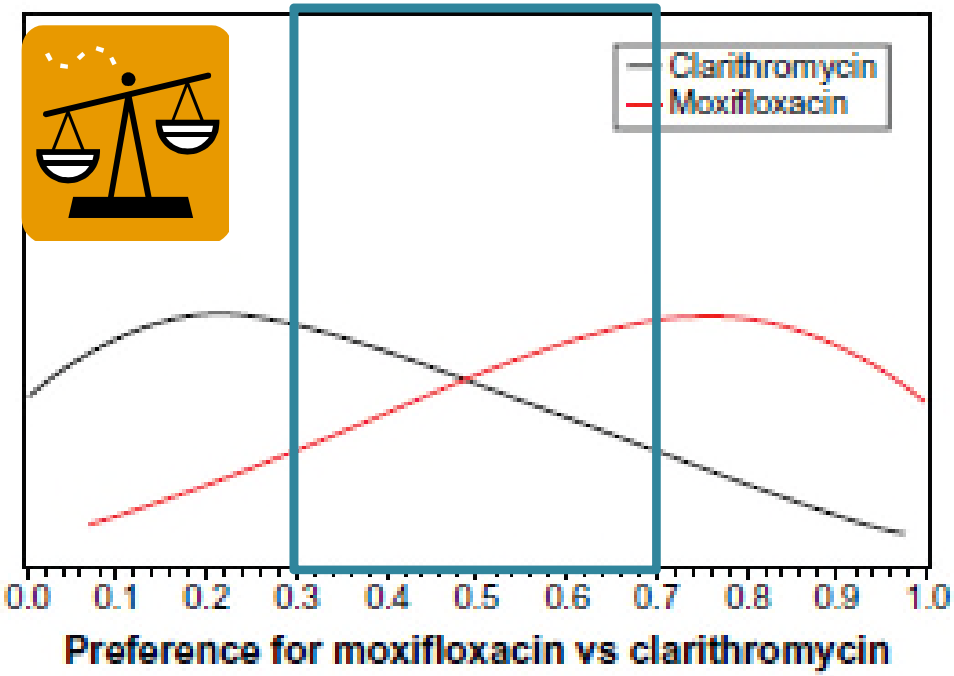
Data Source – Walker et al Example

- Patients from Pennsylvania Pharmacy Assistance Contract for the Elderly (PACE) linked to Medicare Parts A and B from 2000 – 2006
  - Outpatient ICD-9 codes 482.9, 485 or 486 after 270 days continuous enrollment
  - Radiologic exam of chest up to 3 days before or 2 days after dx
  - Dispensing of single antibiotics for CAP within 3 days of dx
  - Excluded HIV dx in 270 days prior to index, insulin in prior 3 mo or any of the antibiotics under study, hospital discharge in prior 3 mo
- Outpatient and inpatient diagnoses, procedure codes and dates of all inpatient/ outpatient services linked to pharmacy dispensing data and PA vital-statistics files
- Small number of optimal treatments for Community Acquired Pneumonia (CAP) in patients without chronic disease predisposing to drug-resistant *Streptococcus pneumoniae* and without recent antibiotic therapy
- Goal : To identify commonly used treatments for CAP and compare them for disparities in outcomes (hospitalization for pneumonia or antibiotics)

Walker AM *et al*. *Comparative Effectiveness Research* 2013; 3: 11-20.

Preference Score Distributions





- Similar preference – overlapping
- At equipoise – preference scores for both are mostly between .0.3 and 0.7

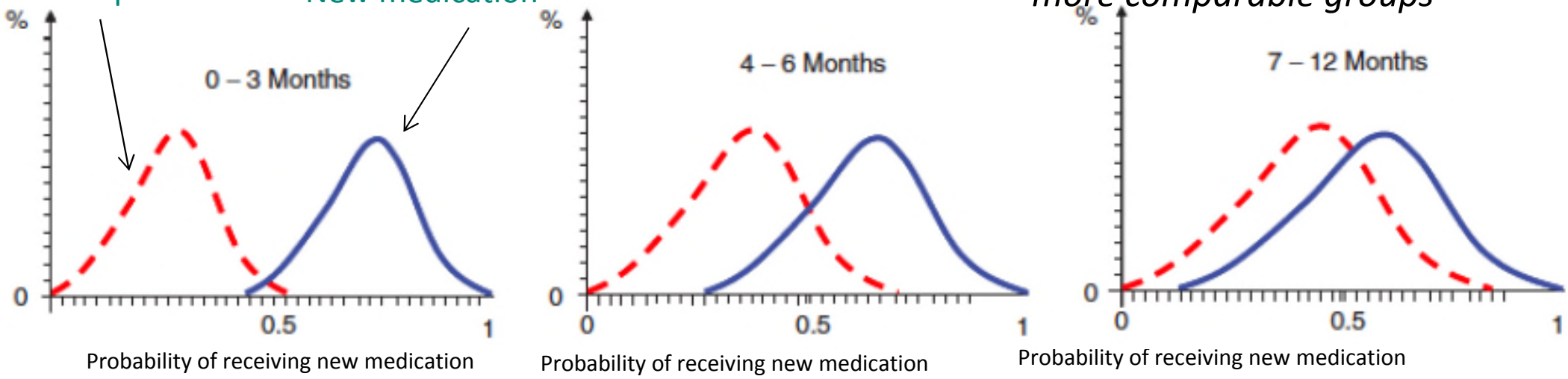
- Differences in preference – mostly not overlapping
- Not at equipoise – preference scores mostly not 0.3 to 0.7

Ref: Walker, JCER 2013

Channeling of Patients to Certain Therapies Often Changes Over Time with Newly Marketed Medications

- When a new medication comes onto the market, sicker patients may receive the new medication at first – over time, patients initiating the new medication may look more like patients initiating a comparator

Increased overlap indicates more comparable groups



→ Increasing time since initial marketing of new medication →

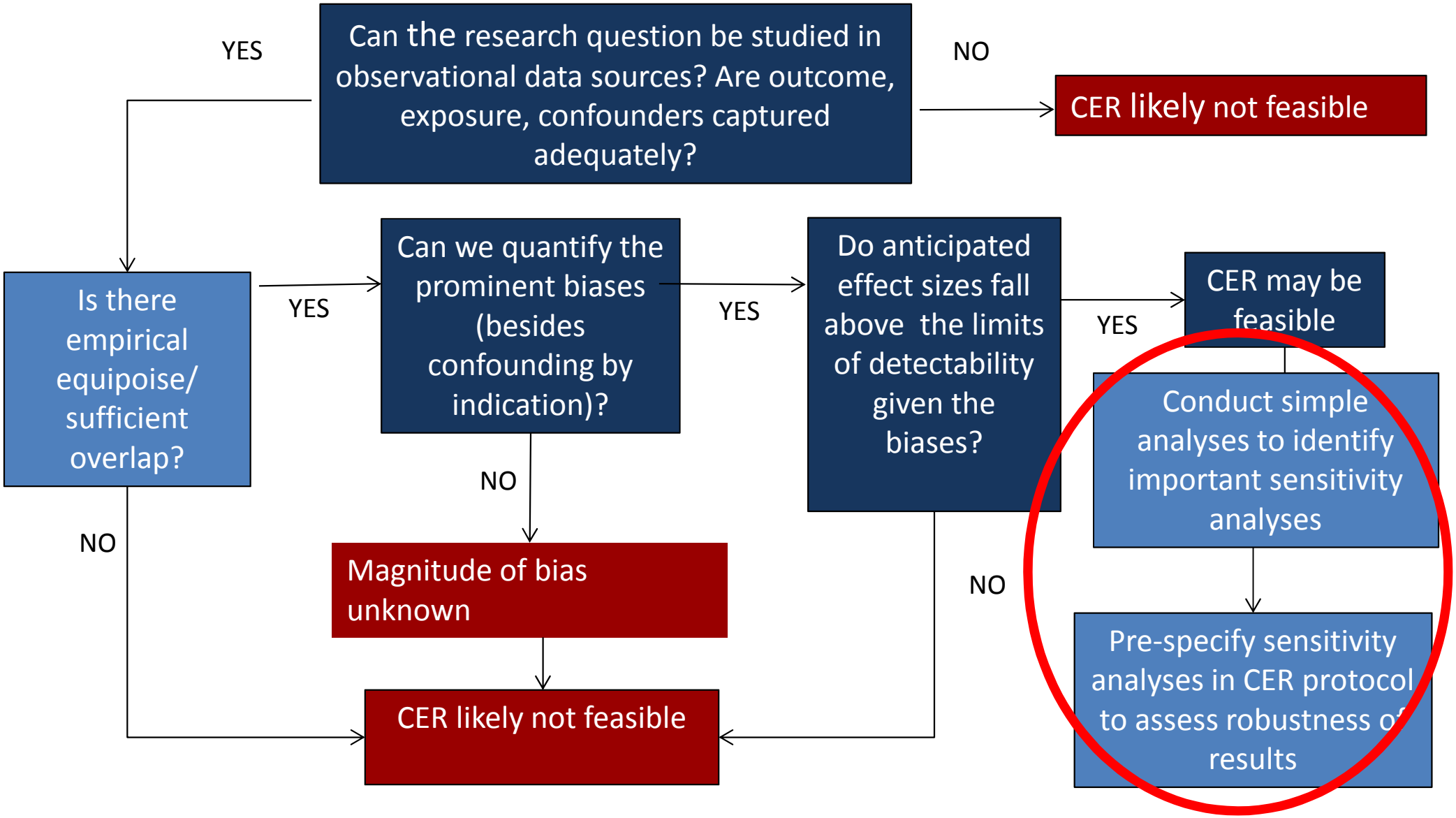
Ref: Schneeweiss et al. *Clin Pharmacol & Ther* 2011; 90: 777-790

What if Empirical Equipoise Does Not Hold?

- Some researchers try to conduct CER anyway, using different methods to control for confounding, and hope for the best
  - Not very interpretable
- Others restrict to only the patients with overlapping PS
  - Limits in generalizability may make it uninterpretable
- If confounding by indication changes over time, some may exclude earlier time periods post-launch, or use time-varying methods
- Some may decide that randomization is needed to balance the groups, and enhance the validity and interpretability of the results

" I miss the good old days ! "

Steps for Assessing Feasibility and Important Sensitivity Analyses to Pre-Specify Prior to CER Launch



Girman CJ, Faries D, Ryan P, Rotelli M, Belger M, Binkowitz B, O'Neill R. *J Comparative Effectiveness Research* 2014; 3: 259-270.
Walker AM *et al*. *Comparative Effectiveness Research* 2013; 3: 11-20

SENSITIVITY ANALYSES

- Assume you are planning a CER study in a claims database
- In drafting a protocol, there are a number of design parameters and definitions/algorithms that must be applied
- Many these decisions about design parameters and definitions could impact on your results
- How do you decide what sensitivity analyses should be performed?
  - Could assess robustness of results to variations in every possible definition and assumption
    - Makes interpretation difficult
    - Greatly raises risk of spurious findings
  - Could select only those your team believes are the ‘top’ ones based on subjective judgment

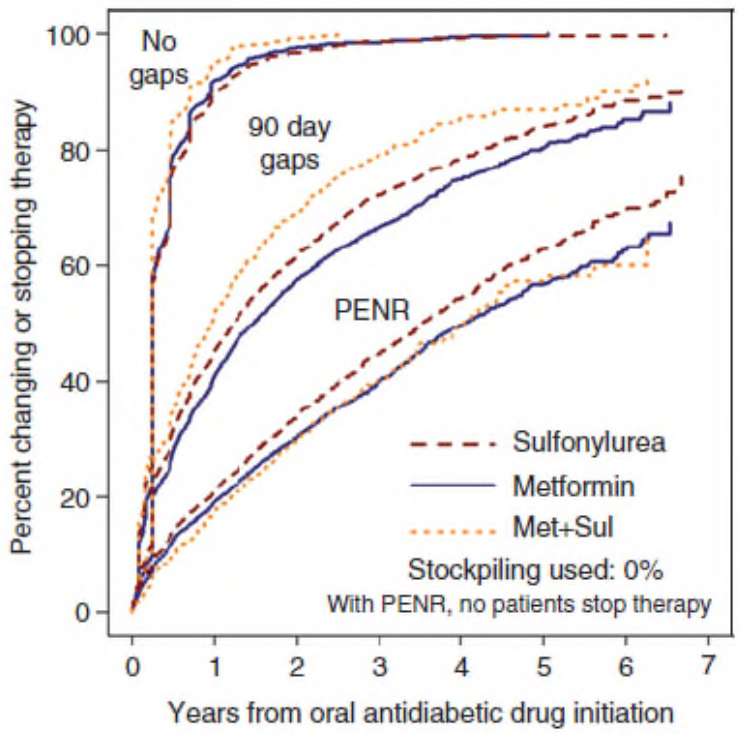
SENSITIVITY ANALYSES

- The number and types of design parameters that could influence results is extensive
  - Current guidelines consider sensitivity analyses mostly to address statistical assumptions
  - Few recommend testing robustness of results to variations in definitions and algorithms to define outcome, exposure, baseline period and handle of irregularities in exposure
  - Interpretation of CER in the context of extensive sensitivity analyses could be daunting, and result in spurious findings
  - Regardless, such sensitivity analyses should be pre-specified as part of the protocol (AHRQ Users Guide)

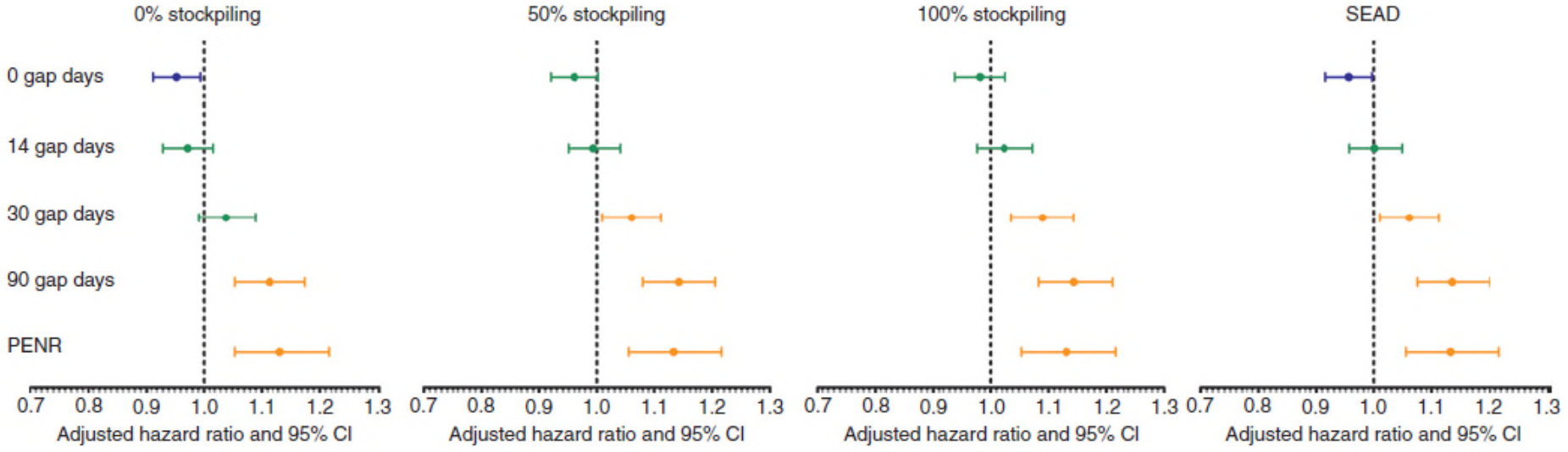
Need a rationale approach to identify *a priori* which sensitivity analyses are most important

Schneeweiss S. *Pharmacoepidemiology and Drug Safety* 2006; 15: 291-303.
Greevy RA *et al*. *Clin Pharmacol Ther* 2011;90:813-819.

Illustration of Impact of Design Decisions



- Impact of different definitions and rules for handling exposure 'gaps' on results
- Statistically significant effects in both directions depending on gap handling
- Suggests much more attention should be paid to such decisions in CER



Greevy et al. Clin Pharmacol Ther 2011; 90: 813-819

POTENTIAL SENSITIVITY ANALYSES

Design Feature	Relevance	Sensitivity Analysis
Look-back window **Baseline period before therapy - often 12 mo, sometimes 6**	Potential misclassification of the outcome, 'new user', baseline confounders with shorter look-back. Recent research proposed using all available data.	Apply different look-back windows. Lose N with longer windows. Require medical encounter during look-back to better capture comorbidities
Outcome definitions / algorithms based on medical dx and/or CPT codes	Potential misclassification of outcome if too narrow or too broad. Coding of diagnoses can vary widely and be subject to bias.	Apply broader / more narrow definitions. Chart review often needed to validate cases, at least in a subset.

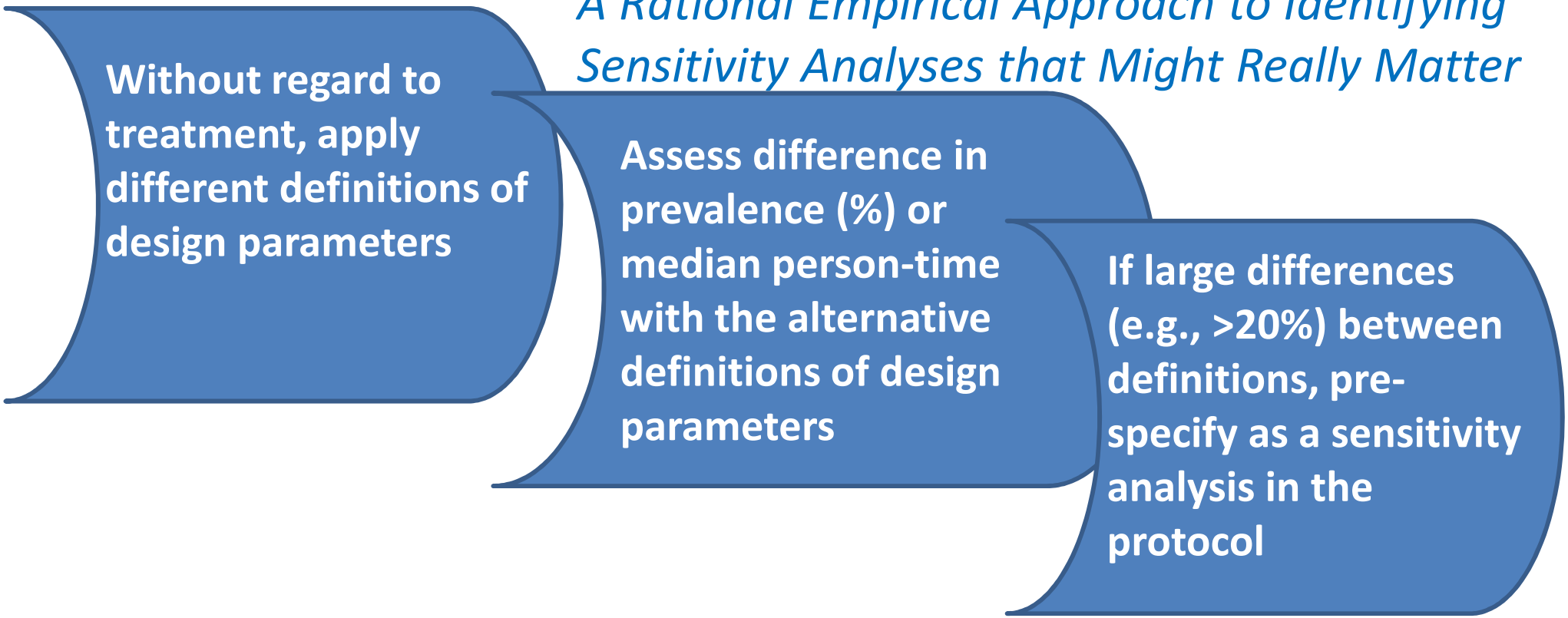
Girman CJ, Faries D, Ryan P, Rotelli M, Belger M, Binkowitz B, O'Neill R. Pre-Study Feasibility and Identifying Sensitivity Analyses for Protocol Pre-Specification in Comparative Effectiveness Research. *J Comparative Effectiveness Research* (2014; in press).

POTENTIAL SENSITIVITY ANALYSES OF EXPOSURE

Exposure definitions typically based on Rx (prescription) order or dispensing	Exposure misclassified if patients do not fill or take Rx. Data may be missing (OTC meds/$4 generics, paid out of pocket), inaccurate or inconsistent. 'Grace period' allows varying timing of refills	Apply different definitions of exposure; require >=2 Rx to improve likelihood that patients took medication (be careful of other biases (e.g., immortal time bias).
Changes in therapy Gaps in coverage, overlap - switching or augmentation of therapies, discontinuations	Change in therapy after index date often ignored / censored; rarely accounted for as time-varying covariates, even if related to outcome.	Apply different approaches to handling gaps, switching and augmentation of Rx, including accounting for time-varying covariates

Steps for Assessing Which Sensitivity Analyses Should be Pre-Specified

A Rational Empirical Approach to Identifying Sensitivity Analyses that Might Really Matter



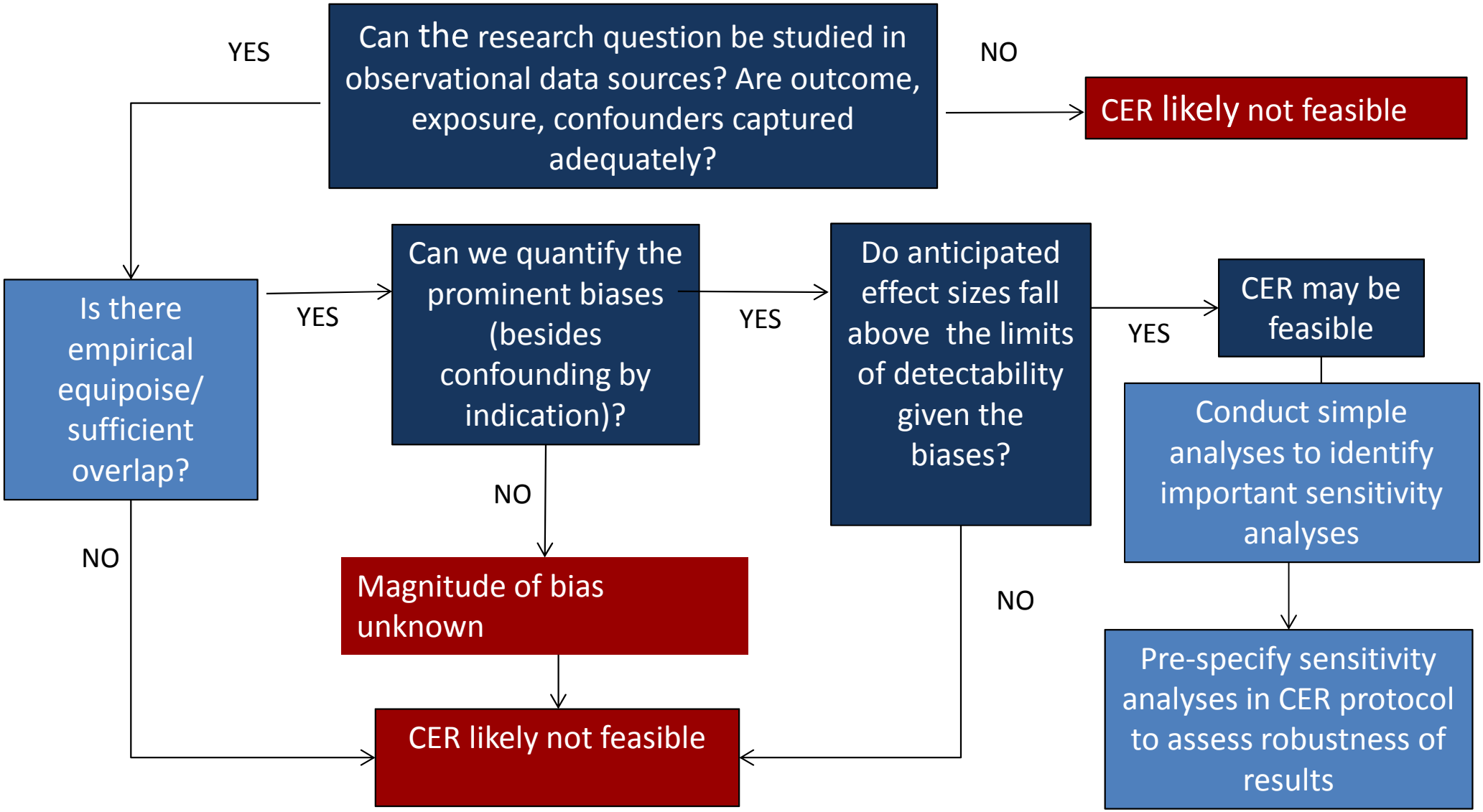
Example: Baseline prevalence of comorbidities defined with a 12 month fixed time window prior to index date vs using all available data

Example: Exposure median person-time with different approaches to handling prescription gaps

Sample Pre-Study Analyses to Identify Important Sensitivity Analyses

DESIGN FEATURE	RELEVANCE	PRE-STUDY ANALYSES (COMBINING TREATMENTS)
Look-back window before therapy	Potential misclassification of outcome, new exposure, baseline confounders	Prevalence of confounders, % prior outcome history, % new users for different look-back windows
Outcome definitions	Potential misclassification of outcome	Incidence differences with broader or more narrow medical diagnosis code definitions of outcomes
Exposure based on Rx order / dispensing	Potential misclassification of exposure	Differences in median person-time with different exposure definitions
Changes in therapy, gaps & augmentation	Potential misclassification of exposure	Differences in median person-time with varying rules for handling gaps, switching, augmentation

Analyses to Pre-Specify Prior to CER Launch



Girman CJ, Faries D, Ryan P, Rotelli M, Belger M, Binkowitz B, O'Neill R. Pre-Study Feasibility and Identifying Sensitivity Analyses for Protocol Pre-Specification in Comparative Effectiveness Research. *J Comparative Effectiveness Research* 2014; 3: 259-270.

DOI: 10.2217/CER.14.16

Special Report

Pre-study feasibility and identifying sensitivity analyses for protocol pre-specification in comparative effectiveness research

Cynthia J Girman[*,1], Douglas Faries[2], Patrick Ryan[3], Matt Rotelli[4], Mark Belger[2], Bruce Binkowitz[5] & Robert O'Neill for the Drug Information Association CER Scientific Working Group

Journal of Comparative Effectiveness Research 2014; 3: 259-270

 www.glasbergen.com

Design Office

"The committee met to approve your idea. But first we had to approve the approval, providing everyone agreed to disagree to approve the agreement which approved the approval agreement. After that, things got complicated."

Thank you

Cynthia J Girman

Executive Director and Head

Data Analytics & Observational Methods

Center for Observational & Realworld Evidence

Merck & Co., Inc

cindy_girman@merck.com