Clinical Trial Eligibility Criteria and the Burden of Generalizability

a call to the OHDSI community

Amelia J Averitt, MPH MA MPhil; Chunhua Weng, PhD; Adler J Perotte, MD MA

Columbia University, Department of Biomedical Informatics

RESEARCH AND PRACTICE

What assumptions do we make when we apply clinical knowledge to the treatment of patients?

Evidence Based Medicine

EVIDENCE BASED MEDICINE

"...a systematic approach to analyze published research as the basis of clinical decision making..." - McMasters University, 1990s[1]

"...the conscientious and judicious use of current best evidence from clinical care research in the management of individual patients..." - Sackett, et al. 1996[2]

EVIDENCE BASED MEDICINE

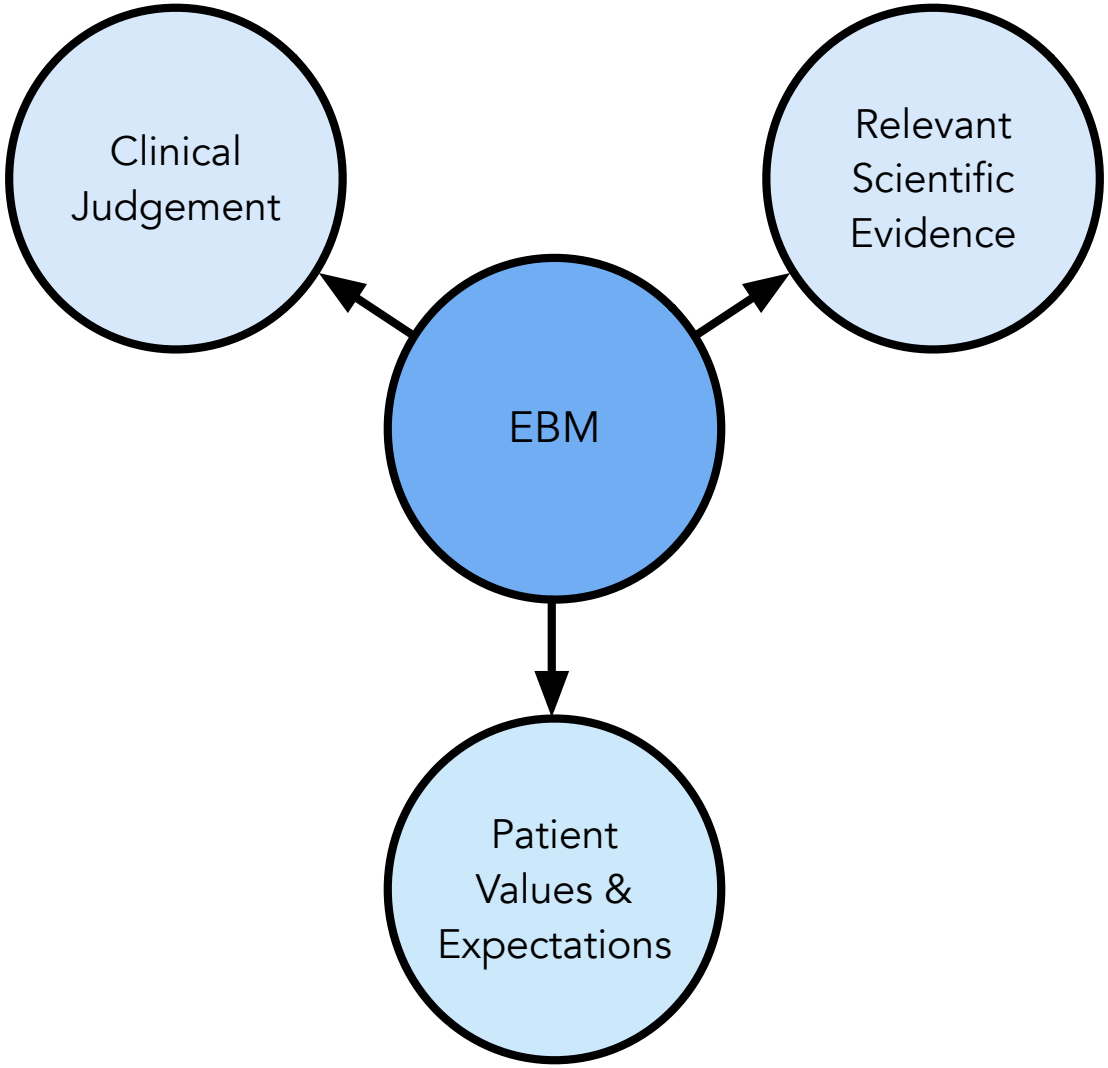


Image adapted from citation 2

THE HIERARCHY OF EVIDENCE BASED MEDICINE

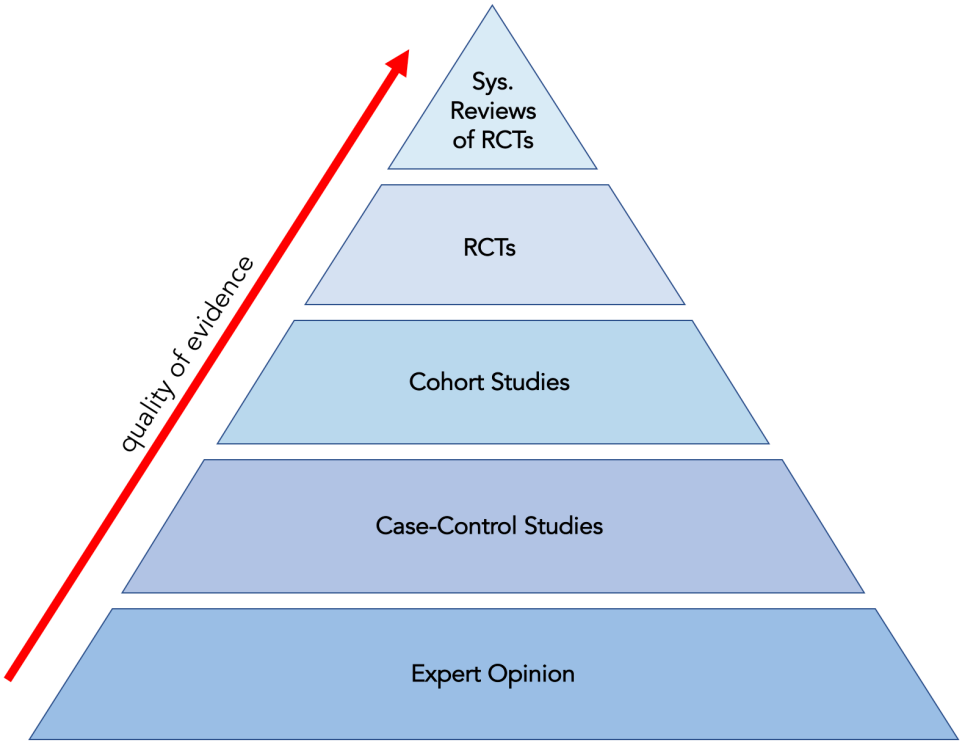


Image adapted from citations 2, 3

"Because the [randomized trial is] so much more likely to inform clinicians and so much **less likely to mislead** them, it has become the "gold standard" for judging whether a treatment does more good than harm." [3]

RCTs are imperfect

factors that ensure high **Internal Validity** ...

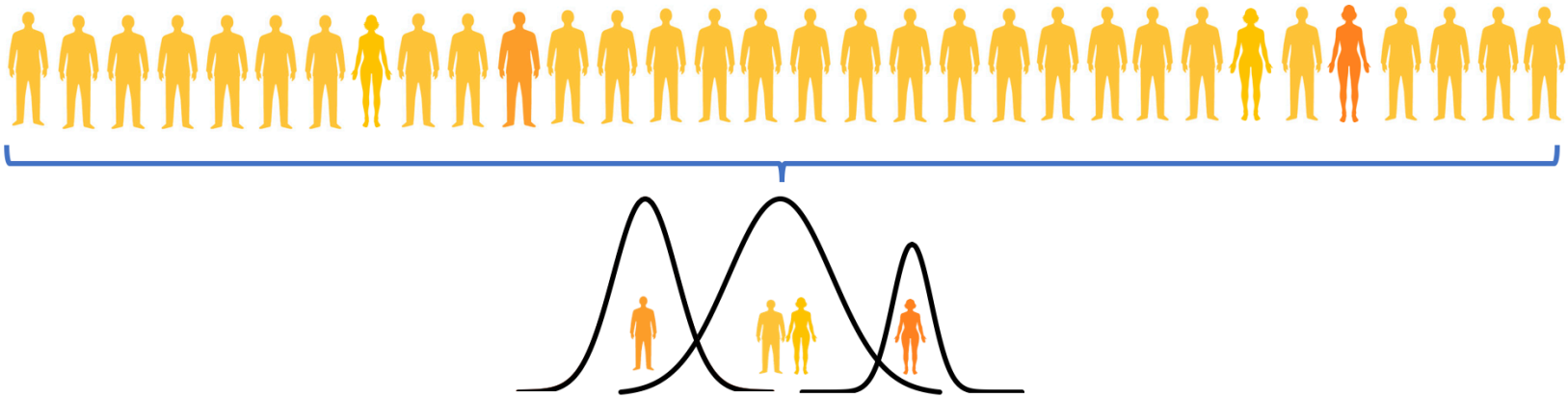
- randomization
- presence of a control
- curated population
- blinding & masking

...may impede **External Validity**. RCTs cited as discriminatory to

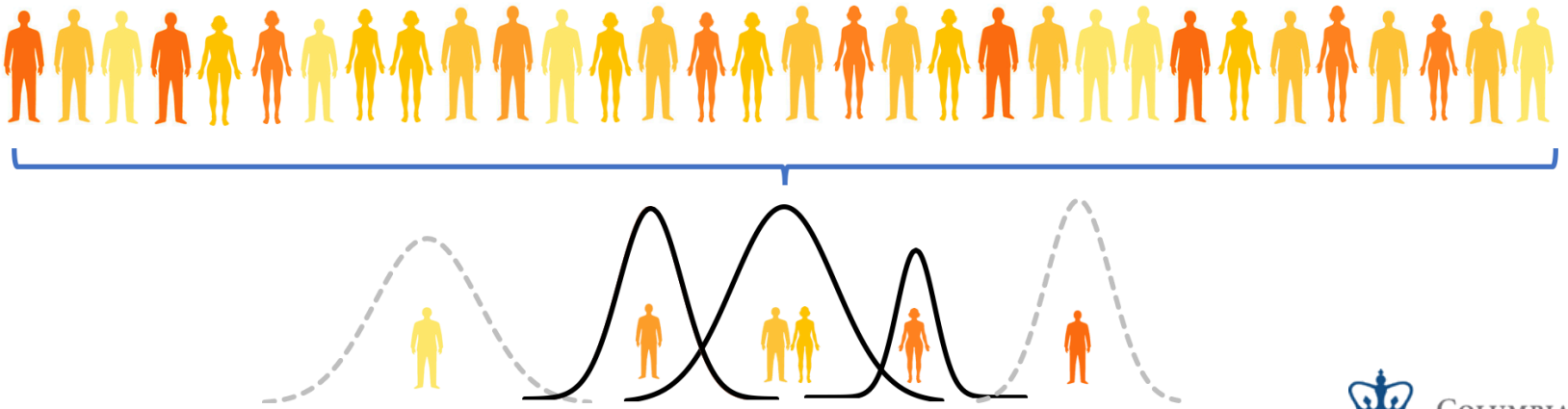
- women [4]
- those with comorbidities [5]
- the elderly [4,6]
- minorities [4,7]

RCT EFFECT ESTIMATES AS A WEIGHTED AVERAGE

Different patient types may demonstrate different responses, known as **heterogeneity of treatment effect.** Reported RCT effect estimates are a function of types of patients that were studied.



In the real world, if the distribution of patient types changes, we cant reliably expect replication of the RCT effect estimate.



Amelia J. Averitt, MPH MA MPhil

Columbia University
Department of Biomedical Informatics

THE PROBLEM

EBM advocates that clinicians apply RCT-generated knowledge to patients. But this assumes that the treated population outside of the study the the same as the RCT population.

We know that RCTs study a small, often homogeneous subpopulation, that is likely not representative.

Therefore, can't assume that trial's distribution is the same as the target population patient and that we will see same effect estimate.

So, how do we identify the applicable patients?

ELIGIBILITY CRITERIA

RCT eligibility criteria should identify applicable patients, for which the effect estimate replicates.

Image credit for hand: Jamie Yeo, The Noun Project

ELIGIBILITY CRITERIA

RCT eligibility criteria should identify applicable patients, for which the effect estimate replicates.

... But do they?

Image credit for hand: Jamie Yeo, The Noun Project

THE GOAL & EARLY STUDIES

We want to better understand the roll of RCTs in evidence based medicine. As a preliminary step, we conducted two trials:

Study 1: Explore the impact of eligibility criteria on the effect estimate.

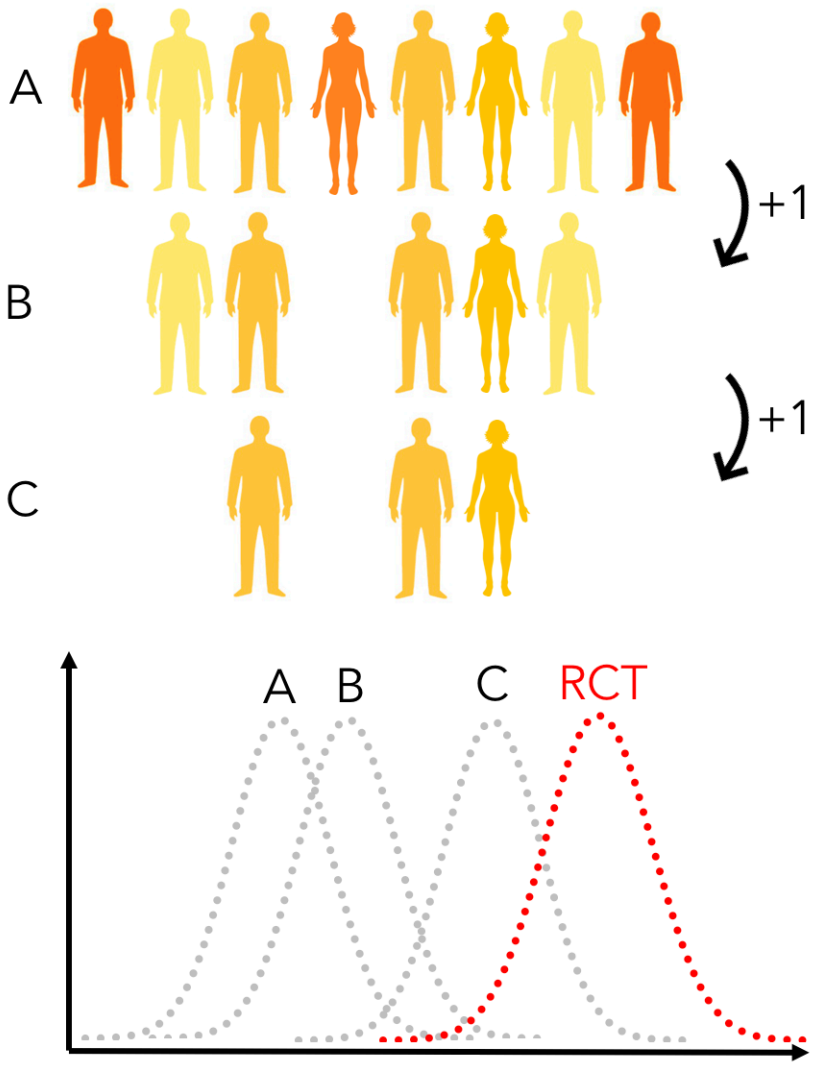
- does heterogeneity of treatment effect exist?

Study 2: Examine potential sources of residual bias in effect estimates.

- is there covariate balance between the trial and real-world populations?

STUDY 1

Study 1: Hypothesis



We hypothesize that the incremental addition of RCT eligibility criteria to an observational cohort will bring the observational effect estimate closer to the RCT effect estimate.

Image credit for arrow: Star and Anchor Design, The Noun Project

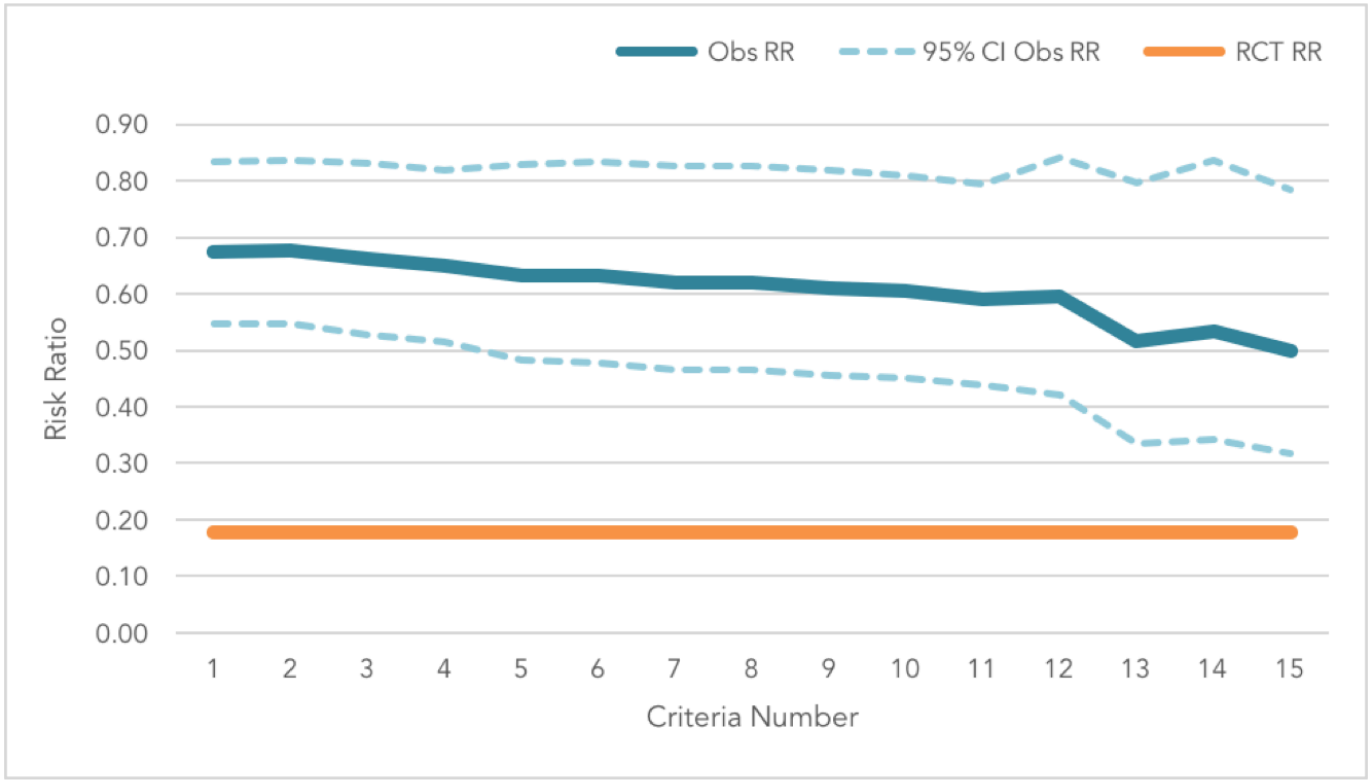
Study 1: Methods

1. Construct a baseline study population using Columbia University Medical Center EHR data according to RCT indication.
2. Incrementally add inclusion and exclusion criteria to baseline cohort using OHDSI analytic tools.
3. Examine the impact of eligibility criteria on the effect estimate.

Efficacy and Tolerability of Sitagliptin Compared with Glimepiride in Elderly Patients with Type 2 Diabetes Mellitus and Inadequate Glycemic Control: A Randomized, Double-Blind, Non-Inferiority Trial

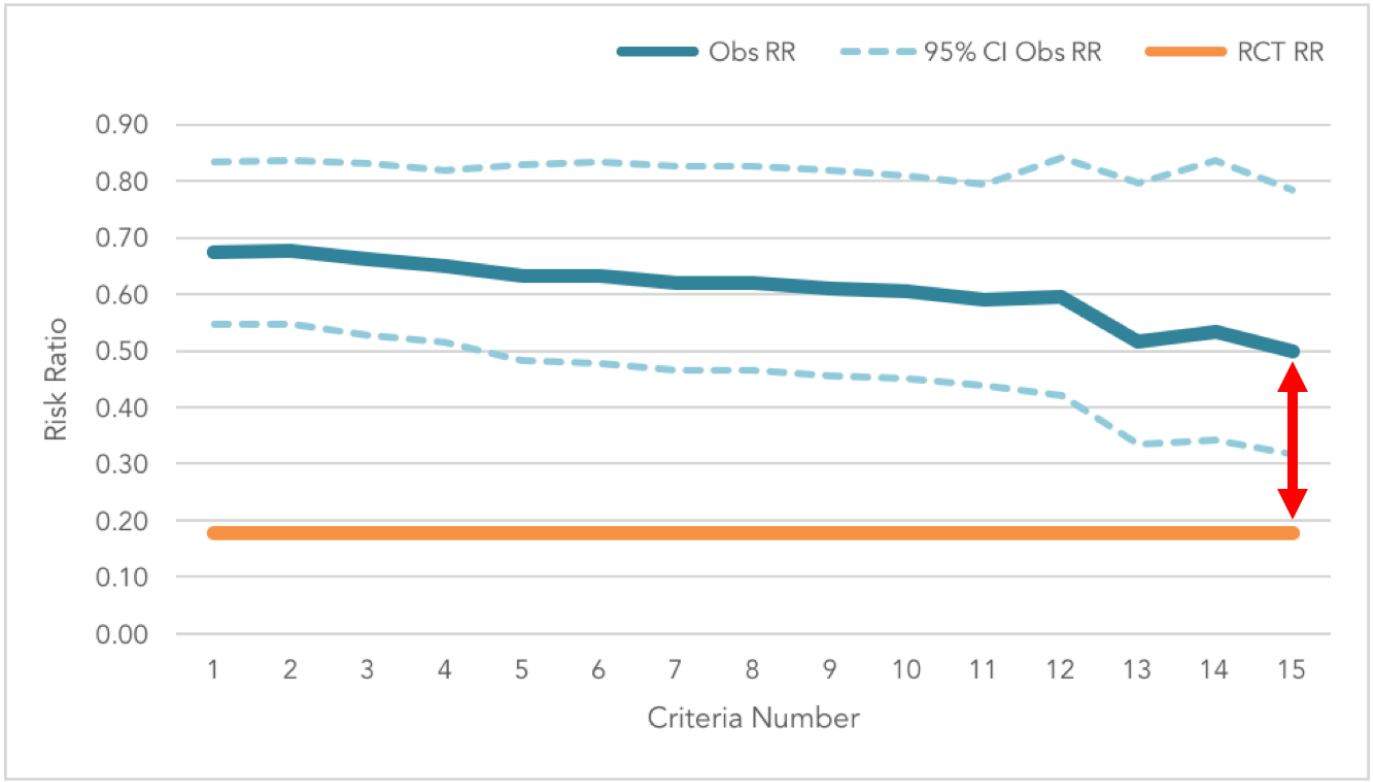
Paul Hartley[1] · Yue Shentu[2] · Patricia Betz-Schiff[2] · Gregory T. Golm[2] · Christine McCrary Sisk[2] · Samuel S. Engel[2] · R. Ravi Shankar[2]

Study 1: Results, risk ratio for hypoglycemia



1. No eligibility criteria
2. #1 + No HIV
3. #2 + No Type 1 DM
4. #3 + No Surgery
5. #4 + No CVD
6. #5 + No Hep. Dis.
7. #6 + No PVD
8. #7 + No High TGs
9. #8 + No Insulin/GLP-1
10. #9 + No PPAR
11. #10 + No DPP-4
12. #11 + No Cancer
13. #12 + No Heme. Dis.
14. #13 + No GFR < 35
15. #14 + No Hx SA

Study 1: Results, risk ratio for hypoglycemia



1. No eligibility criteria
2. #1 + No HIV
3. #2 + No Type 1 DM
4. #3 + No Surgery
5. #4 + No CVD
6. #5 + No Hep. Dis.
7. #6 + No PVD
8. #7 + No High TGs
9. #8 + No Insulin/GLP-1
10. #9 + No PPAR
11. #10 + No DPP-4
12. #11 + No Cancer
13. #12 + No Heme. Dis.
14. #13 + No GFR < 35
15. #14 + No Hx SA

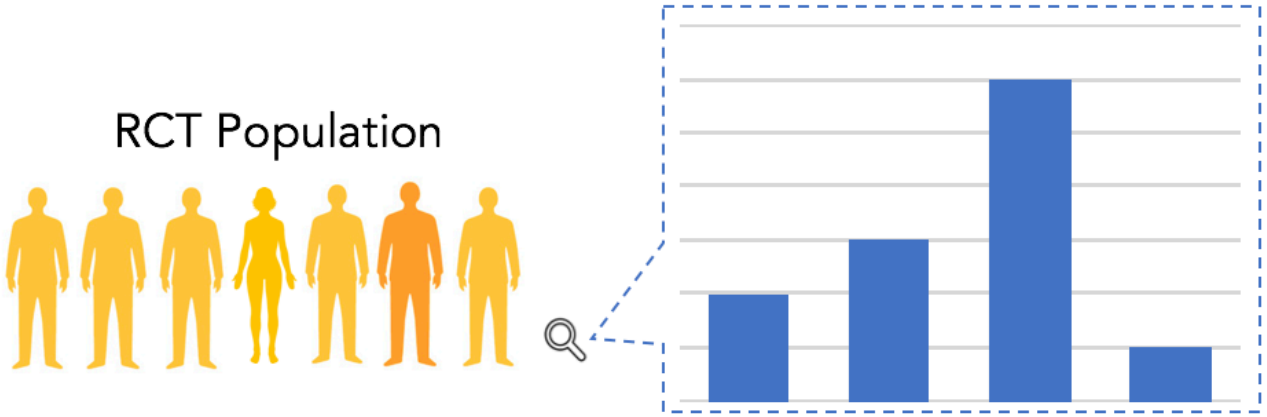
STUDY 2

Study 2: Hypothesis

We hypothesized that the residual bias seen in Study 1 is due to distributional differences in potentially confounding variables.

Study 2: Hypothesis

We hypothesized that the residual bias seen in Study 1 is due to distributional differences in potentially confounding variables.



Study 2: Hypothesis

We hypothesized that the residual bias seen in Study 1 is due to distributional differences in potentially confounding variables.

Amelia J. Averitt, MPH MA MPhil

Columbia University
Department of Biomedical Informatics

Study 2: Methods

1. Observational cohorts were created using Columbia University Medical Center EHR data according to the protocols of three Landmark clinical trials
    - Indication Only
    - Indication + Other Eligibility Criteria
2. Query cohorts to obtain the Table 1 data of their corresponding RCT
3. Compare this observational cohort data to the RCT Table 1 data

Study 2: Trials

EFFECTS OF LOSARTAN ON RENAL AND CARDIOVASCULAR OUTCOMES IN PATIENTS WITH TYPE 2 DIABETES AND NEPHROPATHY

Barry M. Brenner, M.D., Mark E. Cooper, M.D., Ph.D., Dick de Zeeuw, M.D., Ph.D., William F. Keane, M.D., William E. Mitch, M.D., Hans-Henrik Parving, M.D., Giuseppe Remuzzi, M.D., Steven M. Snapinn, Ph.D., Zhonxin Zhang, Ph.D., and Shahnaz Shahinfar, M.D., for the RENAAL Study Investigators*

EFFECTS OF LOSARTAN ON RENAL AND CARDIOVASCULAR OUTCOMES IN PATIENTS WITH TYPE 2 DIABETES AND NEPHROPATHY

Barry M. Brenner, M.D., Mark E. Cooper, M.D., Ph.D., Dick de Zeeuw, M.D., Ph.D., William F. Keane, M.D., William E. Mitch, M.D., Hans-Henrik Parving, M.D., Giuseppe Remuzzi, M.D., Steven M. Snapinn, Ph.D., Zhonxin Zhang, Ph.D., and Shahnaz Shahinfar, M.D., for the RENAAL Study Investigators*

Intensive versus Moderate Lipid Lowering with Statins after Acute Coronary Syndromes

Christopher P. Cannon, M.D., Eugene Braunwald, M.D., Carolyn H. McCabe, B.S., Daniel J. Rader, M.D., Jean L. Rouleau, M.D., Rene Belder, M.D., Steven V. Joyal, M.D., Karen A. Hill, B.A., Marc A. Pfeffer, M.D., Ph.D., and Allan M. Skene, Ph.D., for the Pravastatin or Atorvastatin Evaluation and Infection Therapy–Thrombolysis in Myocardial Infarction 22 Investigators*

Study 2: Trials

EFFECTS OF LOSARTAN ON RENAL AND CARDIOVASCULAR OUTCOMES IN PATIENTS WITH TYPE 2 DIABETES AND NEPHROPATHY

Barry M. Brenner, M.D., Mark E. Cooper, M.D., Ph.D., Dick de Zeeuw, M.D., Ph.D., William F. Keane, M.D., William E. Mitch, M.D., Hans-Henrik Parving, M.D., Giuseppe Remuzzi, M.D., Steven M. Snapinn, Ph.D., Zhonxin Zhang, Ph.D., and Shahnaz Shahinfar, M.D., for the RENAAL Study Investigators*

Intensive versus Moderate Lipid Lowering with Statins after Acute Coronary Syndromes

Christopher P. Cannon, M.D., Eugene Braunwald, M.D., Carolyn H. McCabe, B.S., Daniel J. Rader, M.D., Jean L. Rouleau, M.D., Rene Belder, M.D., Steven V. Joyal, M.D., Karen A. Hill, B.A., Marc A. Pfeffer, M.D., Ph.D., and Allan M. Skene, Ph.D., for the Pravastatin or Atorvastatin Evaluation and Infection Therapy–Thrombolysis in Myocardial Infarction 22 Investigators*

Benazepril plus Amlodipine or Hydrochlorothiazide for Hypertension in High-Risk Patients

Kenneth Jamerson, M.D., Michael A. Weber, M.D., George L. Bakris, M.D., Björn Dahlöf, M.D., Bertram Pitt, M.D., Victor Shi, M.D., Allen Hester, Ph.D., Jitendra Gupte, M.S., Marjorie Gatlin, M.D., and Eric J. Velazquez, M.D., for the ACCOMPLISH trial investigators*

Study 2: some results for RENAAL

	RCT	Columbia University Medical Center		p
		Indication Only	with Eligibility Criteria	H_0: EC = RCT[§]
n =	1,513	3,818	72	
Age	60.00	63.61	59.27	0.390[*]
Gender				<0.002[†]
Male	63.19%	40.78%	38.89%	
Female	36.62%	59.22%	61.11%	
Race/Ethnicity[**]				<0.002[¶]
Asian	16.66%	0.89%	2.78%	
Black	15.20%	14.43%	9.72%	
White	48.65%	0.58%	5.56%	
Hispanic	18.24%	33.76%	41.67%	
Other	1.26%	29.81%	26.39%	
Unknown	-	20.53%	13.89%	
Amputation	8.86%	1.60%	0.00%	0.042[‡]
Neuropathy	51.02%	19.83%	11.11%	<0.002[‡]
Retinopathy	63.71%	5.40%	4.17%	<0.002[‡]
HbA1c	8.54	7.60	8.24	0.298[*]

[*] T-Test; [†] χ^2 Test; [¶] Fisher; [‡] Z Test of proportions; [§] Adjustment by Holm Sequential Correction[8]; ** Normalized

Study 2: some results for RENAAL

	RCT	Columbia University Medical Center		p
		Indication Only	with Eligibility Criteria	H_0: EC = RCT[§]
n =	1,513	3,818	72	
Age	60.00	63.61	59.27	0.390[*]
Gender				**<0.002[†]**
Male	**63.19%**	**40.78%**	**38.89%**	
Female	**36.62%**	**59.22%**	**61.11%**	
Race/Ethnicity[**]				<0.002[¶]
Asian	16.66%	0.89%	2.78%	
Black	15.20%	14.43%	9.72%	
White	48.65%	0.58%	5.56%	
Hispanic	18.24%	33.76%	41.67%	
Other	1.26%	29.81%	26.39%	
Unknown	-	20.53%	13.89%	
Amputation	8.86%	1.60%	0.00%	0.042[‡]
Neuropathy	51.02%	19.83%	11.11%	<0.002[‡]
Retinopathy	63.71%	5.40%	4.17%	<0.002[‡]
HbA1c	8.54	7.60	8.24	0.298[*]

[*] T-Test; [†] χ^2 Test; [¶] Fisher; [‡] Z Test of proportions; [§] Adjustment by Holm Sequential Correction[8]; ** Normalized

Study 2: some results for RENAAL

	RCT	Columbia University Medical Center		p
		Indication Only	with Eligibility Criteria	H_0: EC = RCT[§]
n =	1,513	3,818	72	
Age	60.00	63.61	59.27	0.390[*]
Gender				<0.002[†]
Male	63.19%	40.78%	38.89%	
Female	36.62%	59.22%	61.11%	
Race/Ethnicity[**]				<0.002[¶]
Asian	16.66%	0.89%	2.78%	
Black	15.20%	14.43%	9.72%	
White	48.65%	0.58%	5.56%	
Hispanic	18.24%	33.76%	41.67%	
Other	1.26%	29.81%	26.39%	
Unknown	-	20.53%	13.89%	
Amputation	8.86%	1.60%	0.00%	0.042[‡]
Neuropathy	51.02%	19.83%	11.11%	<0.002[‡]
Retinopathy	63.71%	5.40%	4.17%	<0.002[‡]
HbA1c	**8.54**	**7.60**	**8.24**	**0.298[*]**

[*] T-Test; [†] χ^2 Test; [¶] Fisher; [‡] Z Test of proportions; [§] Adjustment by Holm Sequential Correction[8]; ** Normalized

Study 2: some results for RENAAL

	RCT	Columbia University Medical Center		p
		Indication Only	with Eligibility Criteria	H_0: EC = RCT[§]
n =	1,513	3,818	72	
Age	60.00	63.61	59.27	0.390[*]
Gender				<0.002[†]
Male	63.19%	40.78%	38.89%	
Female	36.62%	59.22%	61.11%	
Race/Ethnicity[**]				<0.002[¶]
Asian	16.66%	0.89%	2.78%	
Black	15.20%	14.43%	9.72%	
White	48.65%	0.58%	5.56%	
Hispanic	18.24%	33.76%	41.67%	
Other	1.26%	29.81%	26.39%	
Unknown	-	20.53%	13.89%	
Amputation	8.86%	1.60%	0.00%	0.042[‡]
Neuropathy	51.02%	19.83%	11.11%	<0.002[‡]
Retinopathy	63.71%	5.40%	4.17%	<0.002[‡]
HbA1c	8.54	7.60	8.24	0.298[*]

[*] T-Test; [†] χ^2 Test; [¶] Fisher; [‡] Z Test of proportions; [§] Adjustment by Holm Sequential Correction[8]; ** Normalized

Study 2: Summary of Results, Gender

Study 2: Summary of Results, Gender

	RENAAL		
	RCT	Ind.	EC
Male	63.19%	40.78%	38.89%
Female	36.62%	59.22%	61.11%

Eligibility Criteria
exacerbates gender bias

Columbia University
Department of Biomedical Informatics

Study 2: Summary of Results, Gender

	RENAAL		
	RCT	Ind.	EC
Male	63.19%	40.78%	38.89%
Female	36.62%	59.22%	61.11%

Eligibility Criteria **exacerbates** gender bias

	PROVE-IT		
	RCT	Ind.	EC
Male	78.11%	45.92%	54.12%
Female	21.89%	54.98%	45.88%

Eligibility Criteria **corrects** gender bias

Columbia University
Department of Biomedical Informatics

Study 2: Summary of Results, Gender

RENAAL

	RCT	Ind.	EC
Male	63.19%	40.78%	38.89%
Female	36.62%	59.22%	61.11%

Eligibility Criteria **exacerbates** gender bias

PROVE-IT

	RCT	Ind.	EC
Male	78.11%	45.92%	54.12%
Female	21.89%	54.98%	45.88%

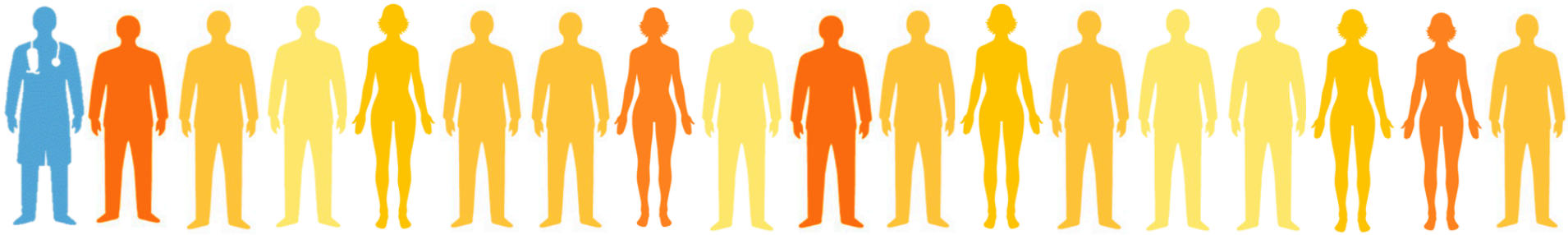
Eligibility Criteria **corrects** gender bias

ACCOMPLISH

	RCT	Ind.	EC
Male	39.48%	67.69%	29.68%
Female	60.52%	32.19%	70.27%

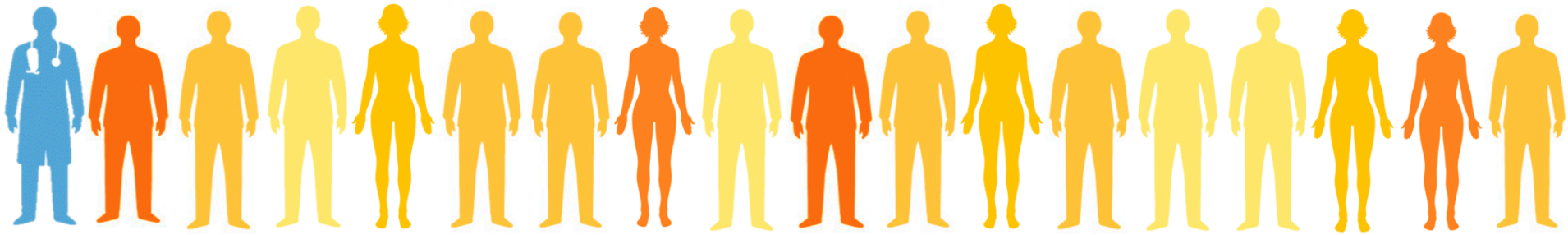
Eligibility Criteria **over-corrects** gender bias

EARLY FINDINGS



EARLY FINDINGS

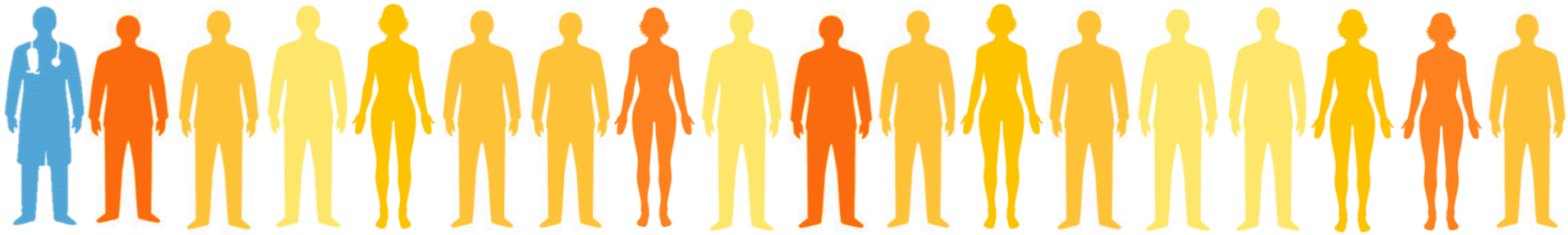
Eligibility criteria are not sufficient to construct a population comparable to the RCT for which the effect estimate will generalize, which suggests a heterogeneity of treatment effect, at least for certain studies.



EARLY FINDINGS

Eligibility criteria are not sufficient to construct a population comparable to the RCT for which the effect estimate will generalize, which suggests a heterogeneity of treatment effect, at least for certain studies.

RCT populations and real-world pops are challenging to compare given fundamental differences.

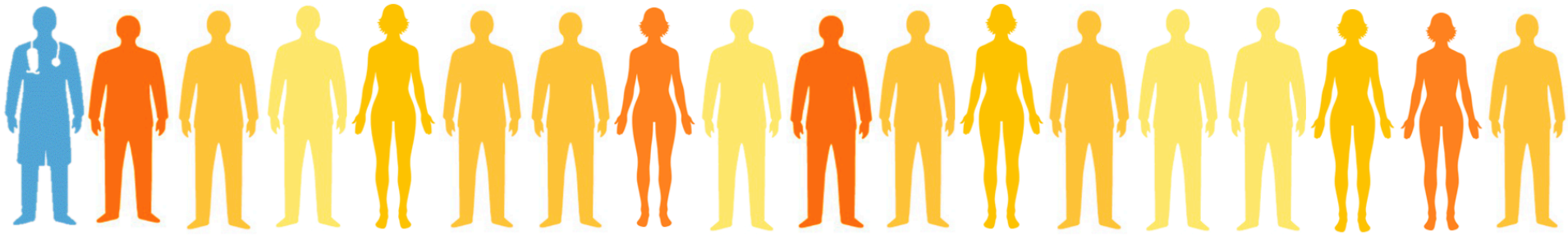


EARLY FINDINGS

Eligibility criteria are not sufficient to construct a population comparable to the RCT for which the effect estimate will generalize, which suggests a heterogeneity of treatment effect, at least for certain studies.

RCT populations and real-world pops are challenging to compare given fundamental differences.

That being said, there is a generalizability problem with our generalizability study **... but you can help!**



HOW YOU CAN HELP

We are looking for collaborators to help us explore RCT applicability, replicability, and generalizability in the context of highly heterogeneous observational data.

You can help in 3 ways.

Ask #1: Replicate the studies presented here today

Do our results generalize to different sites?

The protocol and scripts to replicate these studies with your OMOP CDM-formatted data is available on the OHDSI github:

`OHDSI/StudyProtocolSandbox/Generalizability`

Go run it!

Ask #2: Help us conduct new studies

Do our results generalize to different RCTs?

1. identify a suitable RCT
2. construct a cohort from RCT criteria using ATLAS
3. query your new cohort directly

Share your results and scripts with the community!

Ask #3: Give us your feedback

Go to `forums.ohdsi.org` and comment on our thread.

Generalizability, Applicability, and Replicability of RCTs: A Study

Researchers networkresearch, cdm

References

1. Claridge, J.A. Fabian, T.C. World J. Surg. (2005) 29: 547. https://doi.org/10.1007/s00268-005-7910-1
2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ: British Medical Journal. 1996;312(7023):71-72.
3. Sackett DL. Evidence-based medicine. Semin Perinatol. 1997 Feb;21(1):3–5.
4. Gurwitz, J. H., Col, N. F. Avorn, J. The Exclusion of the Elderly and Women From Clinical Trials in Acute Myocardial Infarction. JAMA J. Am. Med. Assoc. 268, 1417 (1992).
5. Jones, R. Lydeard, S. Prevalence of symptoms of dyspepsia in the community. BMJ 298, 30–2 (1989).
6. Kay, L. Prevalence, incidence and prognosis of gastrointestinal symptoms in a random sample of an elderly population. Age Ageing 23, 146–9 (1994).
7. Bugeja, G., Kumar, A. Banerjee, A. K. Exclusion of elderly people from clinical research: a descriptive study of published reports. BMJ 315, 1059 (1997).
8. A. Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics. 6. 65-70.

INTERESTED? PLEASE REACH OUT!

Amelia J. Averitt
aja2149@cumc.columbia.edu

COLUMBIA UNIVERSITY
Department of Biomedical Informatics