

Testing for equivalence: Establishing benchmarks to validate the use of primary care electronic health records as a chronic disease surveillance tool

Tatem KS, McVeigh KH, Chan PY, Lurie E, Thorpe LE, Perlman SE

Oral Presentation 72450GPmeeting_19

Grant # 72450

Testing for Equivalence:

Establishing benchmarks to validate Primary Care Electronic Health Records as a chronic disease surveillance tool

Kathleen S. Tatem

New York City Department of Health and Mental Hygiene

November 4th, 2015

Presenter Disclosures

Kathleen Tatem

The following personal financial relationships with commercial interests relevant to this presentation existed during the past 12 months:

No relationships to disclose

Overview

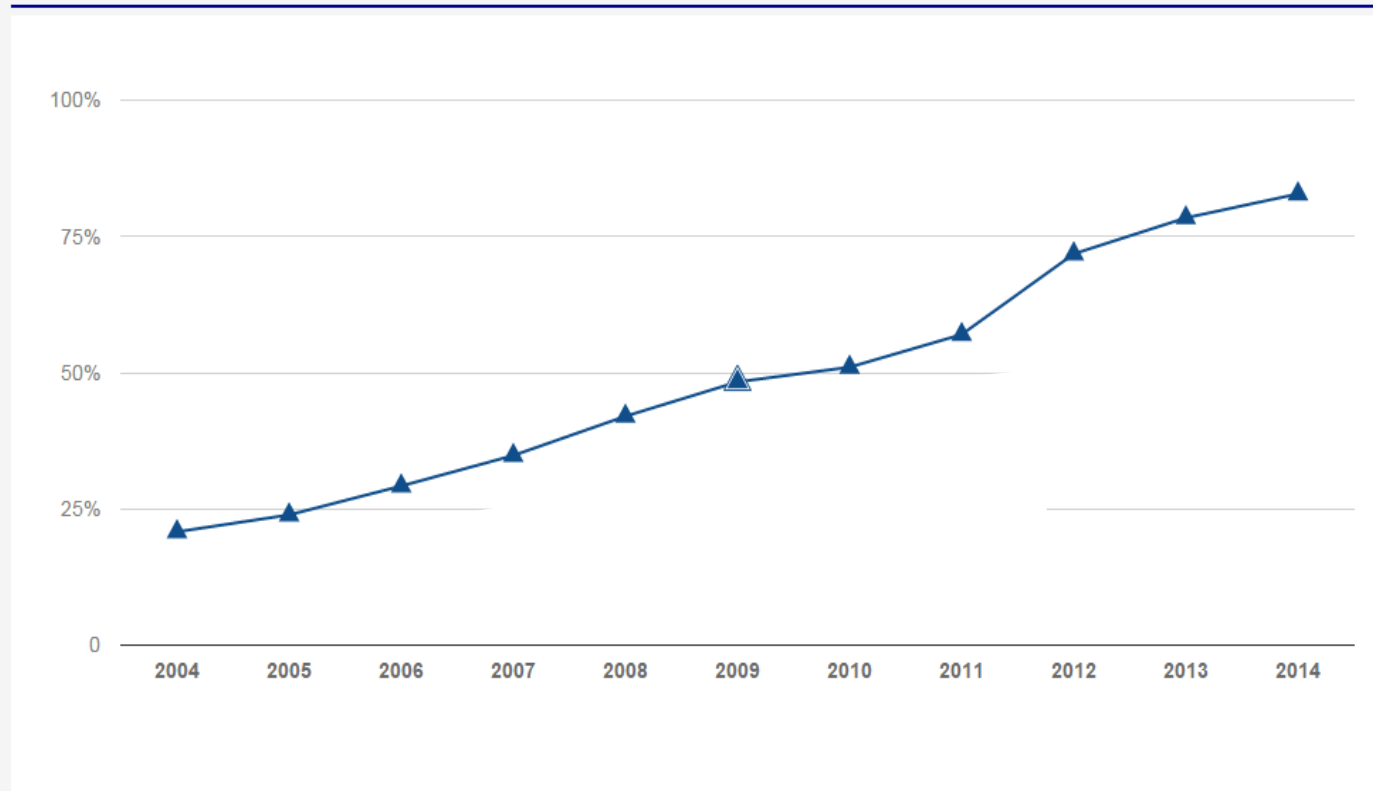
- Background
- Equivalence Testing Methodology
- Benchmarking Study
 - Goals
 - Methods
 - Results
 - Conclusions

Electronic Health Record (EHR) Use

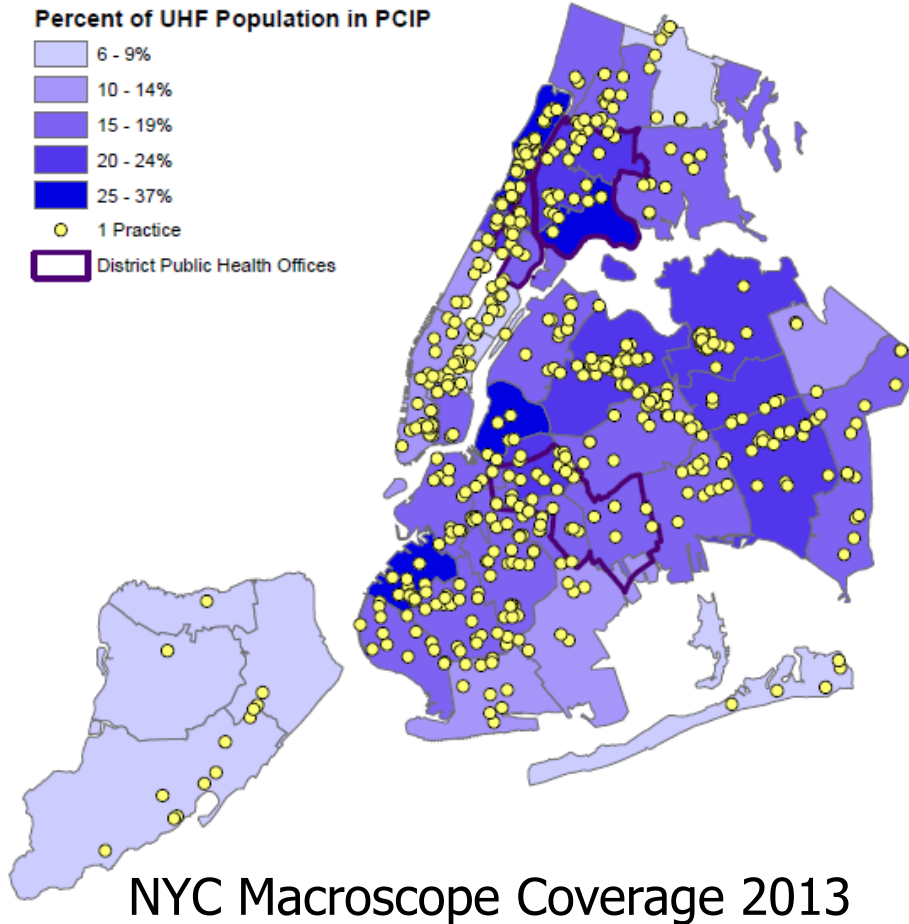
Office-based Physician Electronic Health Record Adoption: 2004-2014

EHR adoption has nearly doubled since 2008

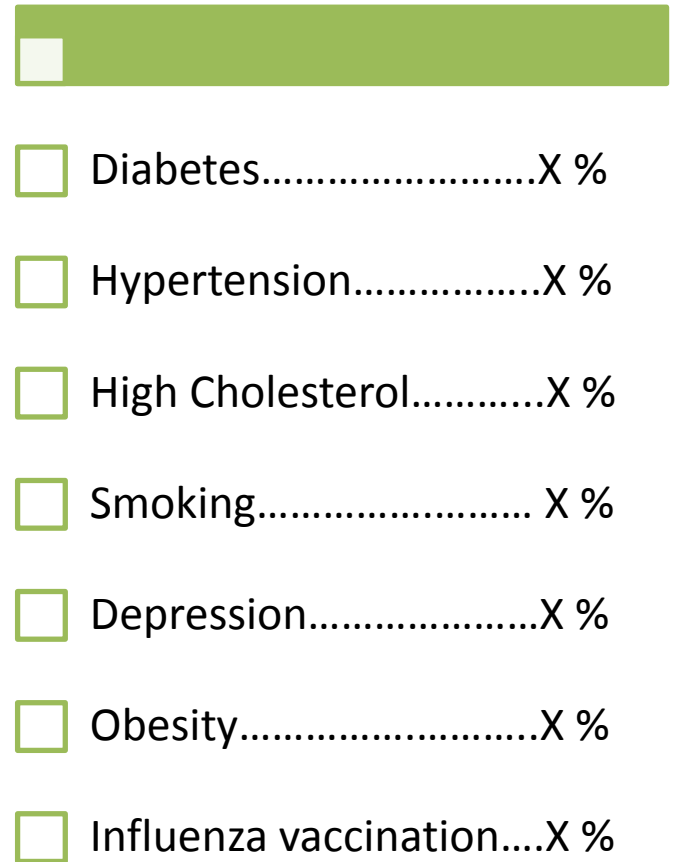
2014

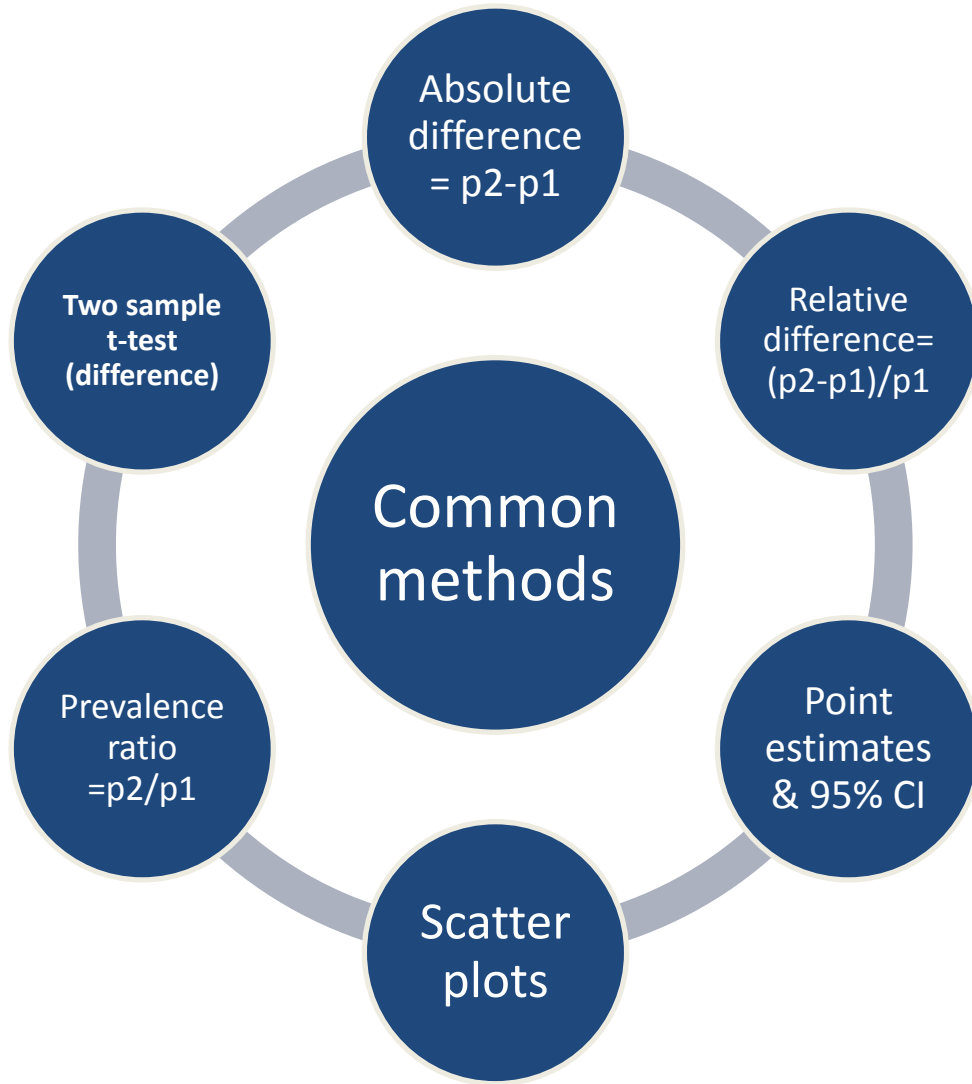


NYC Macroscopic



Chronic Disease Surveillance





Equivalence Testing?

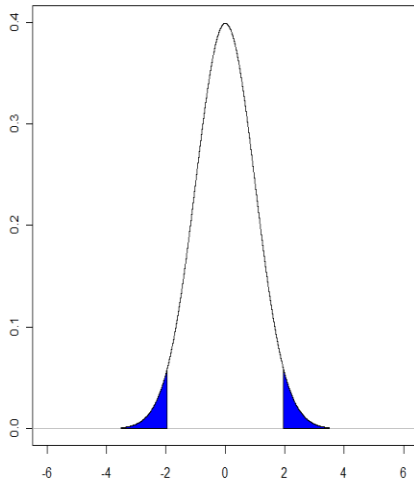


Equivalence testing (TOST)

- **Schuirman's (1987) TOST method:**
 - Conduct **two one-sided t-test** for mean or mean difference
 - Question: “Are these two measures significantly the same?”
 - Commonly used in non-inferiority/equivalence randomized controlled trials
- **KEY:** Establish an “*equivalence margin*” that can be justified

Hypothesis Testing

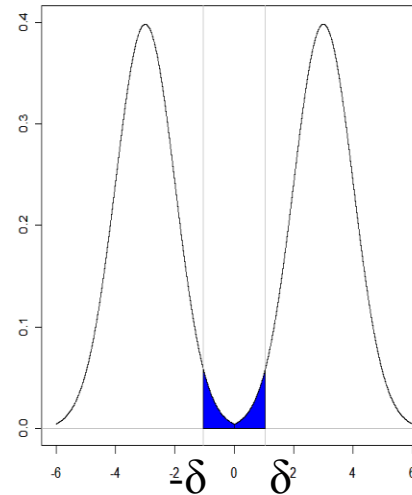
Difference (Two-sided t-test)



Question

“Are these two measures significantly different?”

Equivalence (TOST)



Question

“Are these two measures significantly equivalent?”

KEY

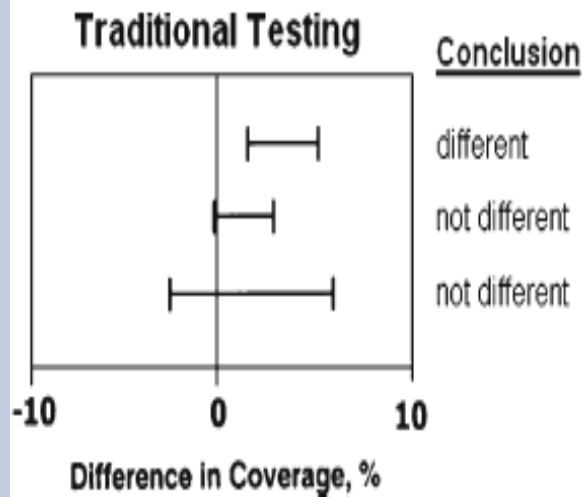
Establish your
“equivalence
margin”
($\pm \delta$)

Images adapted from: <http://www.unt.edu/rss/class/mike/5700/Equivalence%20testing.ppt>

Hypothesis Testing Cont...

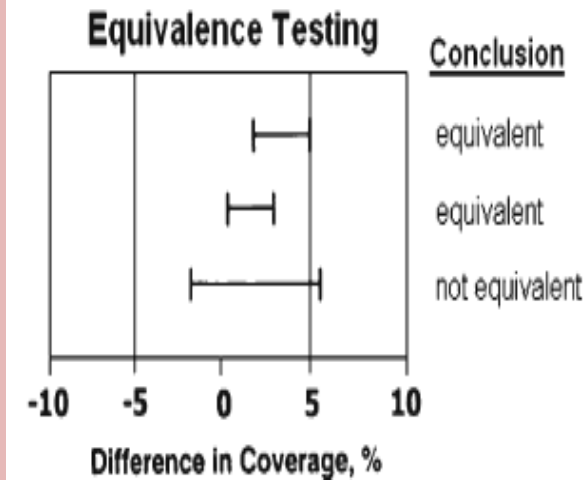
Difference (Two-sided t-test)

- 95% CI = $(1 - \alpha) * 100$



Equivalence (TOST)

- 90% CI = $(1 - 2\alpha) * 100$



Hypothesis Testing Cont...

Equivalence (TOST)

Smoking: TOST Level 0.05 Equivalence Test

Survey	Mean	Lower Bound	90% CL Mean*	Upper Bound	Assessment
Diff (1-2)	-2.8265	-5	> -0.1905 5.4618 <	5	Not equivalent

Test*	Null	DF	t Value	P-Value
Lower	-5	1714.3	1.36	0.0874
Upper	5	1714.3	-4.89	<.0001
Overall				0.0874

*Satterthwaite method for unequal variances

Benchmarking Study Goal

- **Aim 1:** Determine the most appropriate equivalence margins to compare NYC Macroscope prevalence estimates to current gold-standard survey estimates
- **Aim2:** Quantify the maximum difference between NYC HANES and CHS point estimates for health indicators of interest to set empirical benchmark values to provide greater context for interpretation of the NYC Macroscope validation studies.

BENCHMARK STUDY: NYC HANES AND CHS

Methods: Study Populations

- Restricted to “in-care” population ages 20 and older
- Weighted to NYC population based on 2012 and 2013 ACS and age-adjusted to US 2000 Standard Population
 - **NYC HANES (HANES), 2013-2014**
 - Gold-standard survey
 - Household –based sampling with in person interviews, physical examination, collection of biological specimens/labs
 - Estimated N=1135
 - **Community Health Survey (CHS), 2013**
 - Annual, representative, population-based, random-digit dialed telephone survey of adults in NYC modeled—self report
 - Estimated N=6166

*In-care= reported seeing a doctor or health-care professional within the last 12 months

Health Indicators

Diabetes

- Ever told
- *A1C ≥ 6.5 or told + prescribed med (H) vs. Ever told (C)

Hypertension

- Ever told
- *HTN I/II or ever told + prescribed med (H) vs. Ever told (C)

High Cholesterol

- Ever told
- *Chol. ≥ 240 or ever told + prescribed med (H) vs. Ever told (C)

Smoking

- Current Smoker

Depression

- Ever told
- Kessler $6 \geq 13$ (past 30 days, serious psychological distress)

Obesity

- Obese/overweight (BMI ≥ 25)
- Obese (BMI ≥ 30)
- Extreme obesity (BMI > 40)

Influenza vaccination
(within last 12 months)

*Gold standard measure

Methods: Statistical Analysis

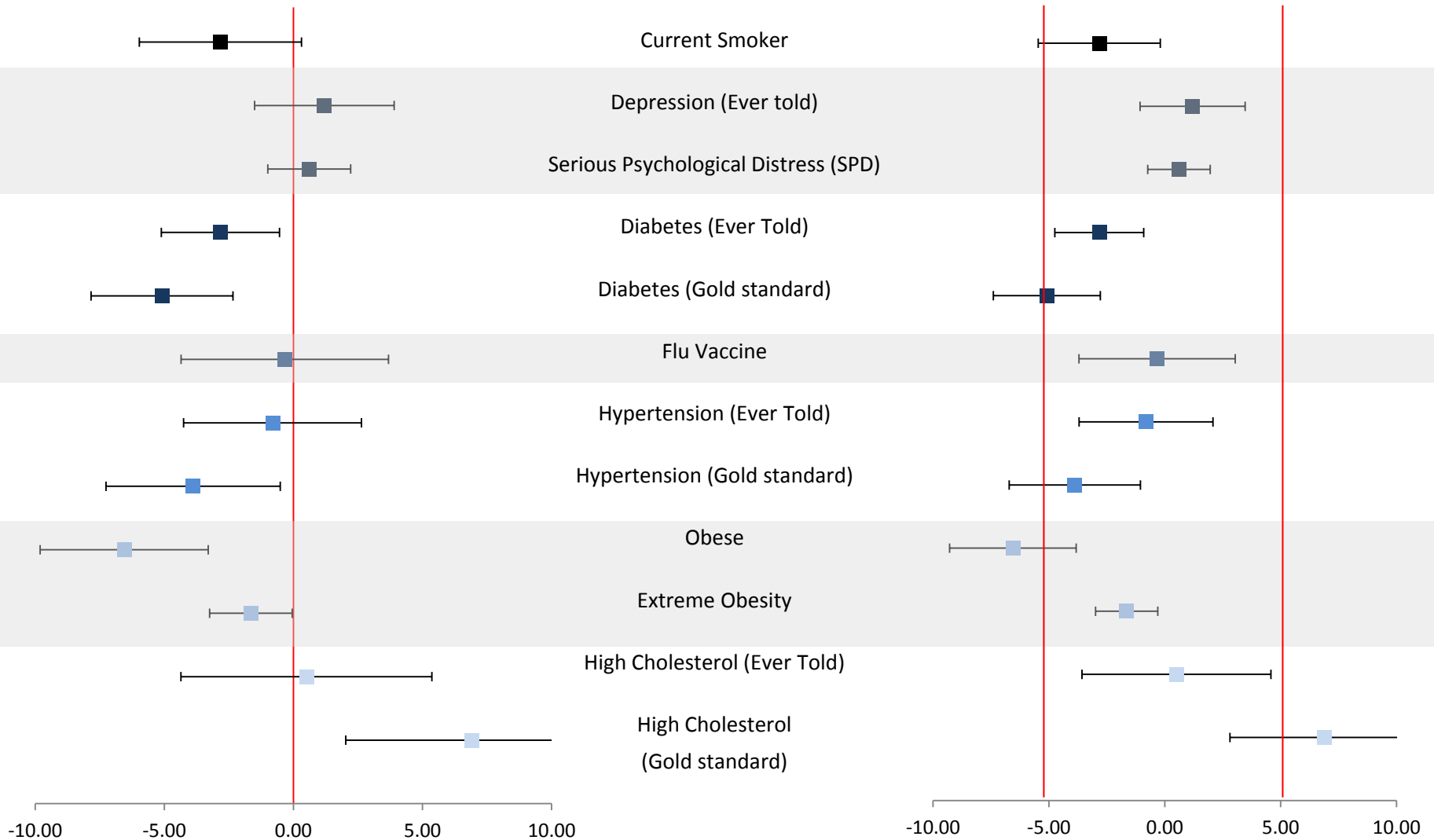
- Prevalence estimates and 95% CI
- **Metrics of goodness-of-fit (*a priori* criteria):**
 - **Equivalence test— TOST**
 - **(-/+ δ = 5; $p < 0.05$)**
 - **Mean difference—two sample t-test ($p < 0.05$)**
 - Prevalence difference (5 points)
 - Prevalence ratio (0.85-1.15)

NYC HANES AND CHS RESULTS

Difference (95% CI)
TTest

CHS vs. NYC HANES

Equivalence (90% CI)
TOST +/-5



CHS vs. NYC HANES (+/-5)

Equivalent and NOT different
(EQUAL)

Not equivalent and not different
(Undetermined)

Equivalent and different
(Equivalent--difference is negligible)

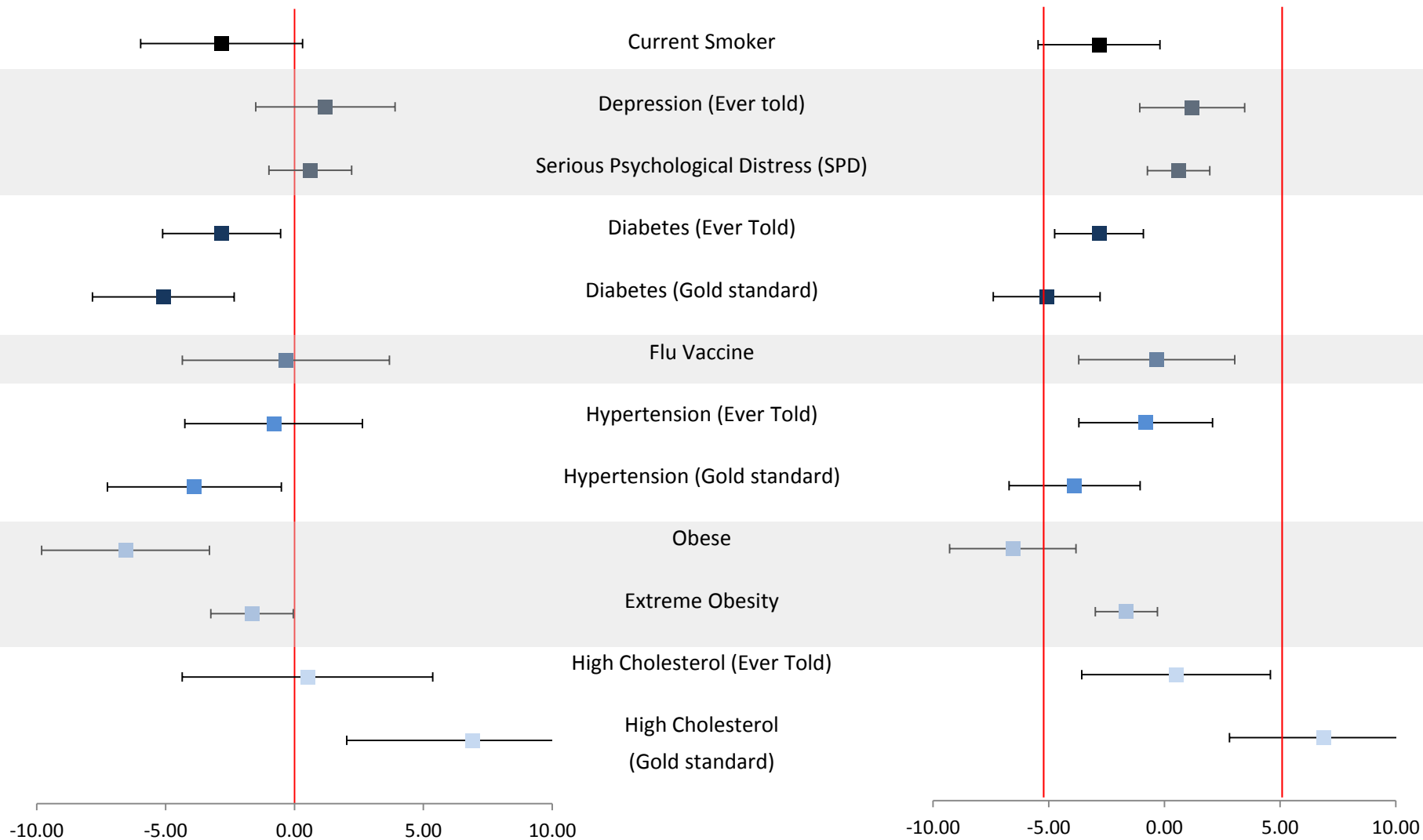
Not equivalent but different
(DIFFERENT)

		EQUIVALENCE TEST	
		Significant (equal)	Not Significant (not equal)
DIFFERENCE TEST	Not Significant (not different)	<ul style="list-style-type: none"> - Depression (Ever told) - Serious Psychological Distress (PSD) - Diabetes (Ever Told) - Flu Vaccine - Hypertension (Ever Told) - High Cholesterol (Ever Told) 	<ul style="list-style-type: none"> - Current Smoker
	Significant (different)	<ul style="list-style-type: none"> - Extreme Obesity 	<ul style="list-style-type: none"> - Diabetes (Gold standard) - Hypertension (Gold standard) - Obese - High Cholesterol— (Gold standard)

Difference (95% CI)
TTest

CHS vs. NYC HANES

Equivalence (90% CI)
TOST (+/-2.5)



CHS vs. NYC HANES (+/-2.5)

Equivalent and NOT different
(EQUAL)

Not equivalent and not different
(Undetermined)

Equivalent and different
(Equivalent--difference is negligible)

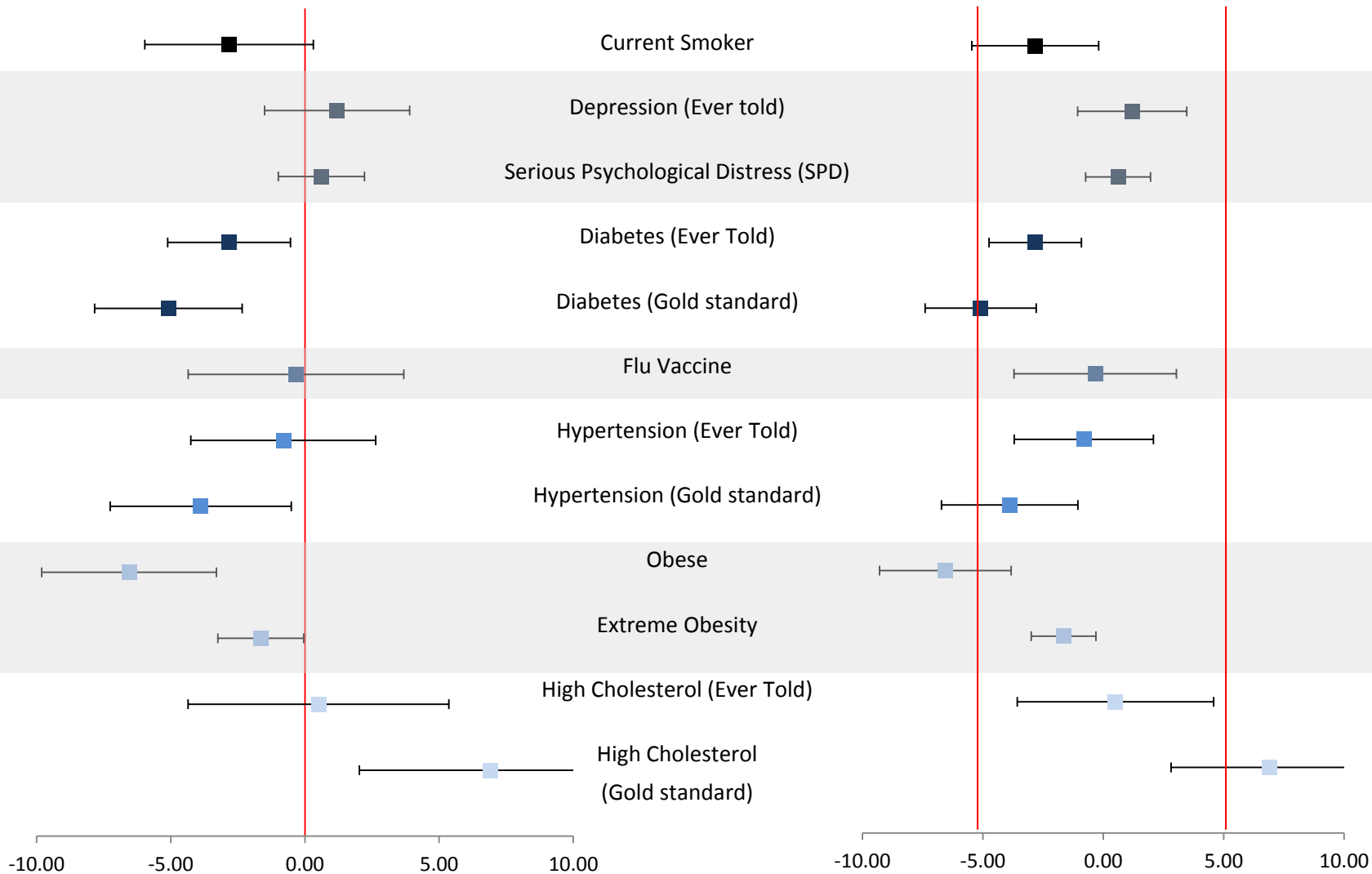
Not equivalent but different
(DIFFERENT)

		EQUIVALENCE TEST	
		Significant (equal)	Not Significant (not equal)
DIFFERENCE TEST	Not Significant (not different)	- Serious Psychological Distress (PSD)	<ul style="list-style-type: none"> - Depression (Ever told) - Current Smoker - Diabetes (Ever Told) - Flu Vaccine - Hypertension (Ever Told) - High Cholesterol (Ever Told)
	Significant (different)		<ul style="list-style-type: none"> - Diabetes (Gold standard) - Hypertension (Gold standard) - Obese - Extreme Obesity - High Cholesterol— (Gold standard)

Difference (95% CI)
TTest

CHS vs. NYC HANES

Equivalence (90% CI)
TOST



CHS vs. NYC HANES (+/-7.5)

Equivalent and NOT different
(EQUAL)

Not equivalent and not different
(Undetermined)

Equivalent and different
(Equivalent--difference is negligible)

Not equivalent but different
(DIFFERENT)

		EQUIVALENCE TEST	
		Significant (equal)	Not Significant (not equal)
DIFFERENCE TEST	Not Significant (not different)	<ul style="list-style-type: none"> - Current Smoker - Depression (Ever told) - Serious Psychological Distress (PSD) - Diabetes (Ever Told) - Flu Vaccine - Hypertension (Ever Told) - High Cholesterol (Ever Told) 	
	Significant (different)	<ul style="list-style-type: none"> - Diabetes (Gold standard) - Hypertension (Gold standard) - Extreme Obesity 	<ul style="list-style-type: none"> - Obese - High Cholesterol— (Gold standard)

Which margin is best?

Margin	Equivalent/ No diff	Not equal/No Diff (Discordant)	Equal/ Diff (Discordant)	Not equal/Different
+/- 2.5	1	5	0	6
+/- 5	6	1	1	4
+/- 7.5	6	0	4	2

Consistently Equivalent: Hypertension(ever told)

CHS % (95% CI)	HANES % (95% CI)	Difference	90% CI (TOST)	95% CI (T-test)
31.6 (30.2-33.0)	32.4 (29.3-35.6)	-0.81	(-3.7,2.1)	(-4.3, 2.6)

	TOST (+/- $\delta=5$)	Student's T-Test
<i>A Priori</i> Criteria	P< .05	P< .05
Benchmark CHS vs. HANES	0.0085	0.6453

Consistently Different: Obesity

CHS % (95% CI)	HANES % (95% CI)		Difference	90% CI (TOST)	95% CI (T-test)
24.7 (23.2-26.3)	31.3 (28.5-34.2)		-6.56	(-9.3,-3.8)	(-9.8,-3.3)

	TOST (+/- $\delta=5$)	Student's T-Test
<i>A Priori</i> Criteria	P < .05	P < .05
Benchmark CHS vs. HANES	0.8256	<.001

Inconclusive: Smoking

CHS % (95% CI)	HANES % (95% CI)	Difference	90% CI (TOST)	95% CI (T-test)
14.9 (13.6-16.3)	17.7 (15.1-20.8)	-2.8	(-5.5,-0.2)	(-6.0, 0.3)

	TOST (+/- $\delta=5$)	Student's T-Test
<i>A Priori</i> Criteria	P< .05	P< .05
Benchmark CHS vs. HANES	0.09	0.08

Exception: Small prevalence, smaller margin?

Extreme Obesity & SPD

Indicator	CHS % (95% CI)	HANES % (95% CI)	Difference	90% CI (TOST)	95% CI (T-test)
Extreme Obesity	3.50 (2.90, 4.21)	5.14 (3.85, 6.83)	-1.7	(-9.3, -3.8)	(-3.3, -0.05)
SPD	5.38 (4.74, 6.09)	4.77 (3.49, 6.47)	0.6	(-0.7, 2.0)	(-1.0, 2.2)

	Indicator	TOST (+/-=5)	TOST (+/-=2.5)	Student's T-Test
<i>A Priori</i> Criteria	—	P < .05	P < .05	P < .05
Benchmark CHS vs. HANES	Extreme Obesity	<.0001	0.1475	0.0438
	SPD	<.0001	0.0107	0.4533

Validation methods

- Metrics of goodness-of-fit (*a priori* criteria):
- Equivalence test— TOST
 - (-/+ = 5; $p < 0.05$)
- Mean difference—two sample t-test ($p < 0.05$)

Conclusions

- Aim1: An equivalence margin of +/- 5 appeared to be the most appropriate margin for estimates with a prevalence estimate larger than 10
 - A margin of +/-2.5 may be a more appropriate margin for indicators with a prevalence estimate less than 10
- Aim 2: Global comparisons of NYC HANES and CHS produced established benchmarks to offer guidance/context to interpret NYC Macroscopic v NYC HANES results

Acknowledgements

DOHMH Division of Epidemiology

Katharine H. McVeigh, PhD, MPH

Elizabeth Lurie, MPH

Claudia Chernov, MPH

Kathleen Tatem

Sharon E. Perlman, MPH

Pui Ying Chan, MPH

Katherine Bartley, PhD

Sungwoo Lim, DrPH

DOHMH Primary Care Information Project

Laura Jacobson, MPH

Lauren Schreiberstein, MA

City University of New York School of Public Health

Lorna Thorpe, PhD

* Special thanks to Jay Bala, Amy Freeman, Ryan Grattan, Carolyn Greene, Tiffany Harris, Stephen Immerwahr, Kevin Konty, Ram Koppaka, Remle Newton-Dame, Jessica Rodriguez-Lopez, Matt Romo, Sarah Shih, Jesse Singer, Elisabeth Snell

This work has been made possible by the financial support of the de Beaumont Foundation, the Robert Wood Johnson Foundation, the Robin Hood Foundation, the NY State Health Foundation, the Doris Duke Charitable Foundation, and the Centers for Disease Control and Prevention

Questions?

Contact:

ktatem@health.nyc.gov

or

nycmacroscope@health.nyc.gov