Disparities by Census Tract-Level Poverty for 53 Reportable Communicable Diseases in New York City

Sharon K. Greene, PhD, MPH,[1]
Alison Levin-Rector, MPH,[1] James L. Hadler, MD, MPH,[2]
Annie D. Fine, MD[1]

[1]Bureau of Communicable Disease
[2]Commissioner's Office

New York City Department of Health and Mental Hygiene

April 16, 2015
CSTE Health Disparities Subcommittee Webinar

Acknowledgments

- All Bureau of Communicable Disease staff

- Additionally, for contributions to interpreting findings for specific diseases:
  - Katherine Bornschlegel, MPH
  - Jennifer Hsieh, MSPH
  - Lucretia Jones, DrPH
  - Sally Slavinski, DVM, MPH

Learning Objectives

- Describe disparities by area-based poverty
  - age-adjusted incidence rate ratios
  - tests for trend
  - population attributable fractions
- List diseases strongly associated with patient residence in high and low poverty areas in NYC
- Develop a process to solicit input for interpreting findings related to many diseases covered by subject matter experts situated across multiple communicable disease surveillance units

Socioeconomic Disparities

- Associated with differential morbidity and mortality for many conditions of public health importance
  - Extent to which this may be true for a wide variety of reportable infectious diseases is unknown
- Area-based poverty measures
  - Characterize socioeconomic conditions of an area (e.g., census tract, ZIP Code)
  - Can be useful for defining populations at increased risk
    - Identifying and tracking disparities
    - Targeting prevention measures

Objective

To describe disparities across area-based poverty levels by systematically analyzing a large set of selected reportable communicable diseases, following the guidance of the Public Health Disparities Geocoding Project[1] as adapted to NYC[2]

1. Krieger N, et al. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: The Public Health Disparities Geocoding Project. *American journal of public health.* Feb 2005;95(2):312-323.

2. Toprani A, Hadler JL. Selecting and applying a standard area-based socioeconomic status measure for public health data: Analysis for New York City. *New York City DOHMH: Epi Research Report, May 2013.*

Inclusion Criteria for Disease Data

- All diseases currently reportable to Bureau of Communicable Disease (BCD), except:
  - Syndromes reflecting a mix of etiologies
    - e.g., encephalitis, bacterial or viral meningitis
  - Rarely reported/not routinely investigated diseases
    - e.g., non-specific *Rickettsia*, hepatitis D
  - Animal diseases
  - (Diseases not reportable to BCD)
    - Tuberculosis, HIV/AIDS, sexually-transmitted infections, some vaccine-preventable infections
  - N=53
- Diagnosed 2006–2013
  - Electronic laboratory reporting mandated in 2006

Defining Neighborhoods

- **Based on census tract**
  - **Smaller-area units preferable to larger-areas units (e.g., ZIP Codes), to increase socioeconomic homogeneity within a unit**
- **Assigned each case to a census tract**
  - **Geocoded using NYC Department of City Planning's Geosupport System[1]**
  - **Determined census tract of residence at time of report**

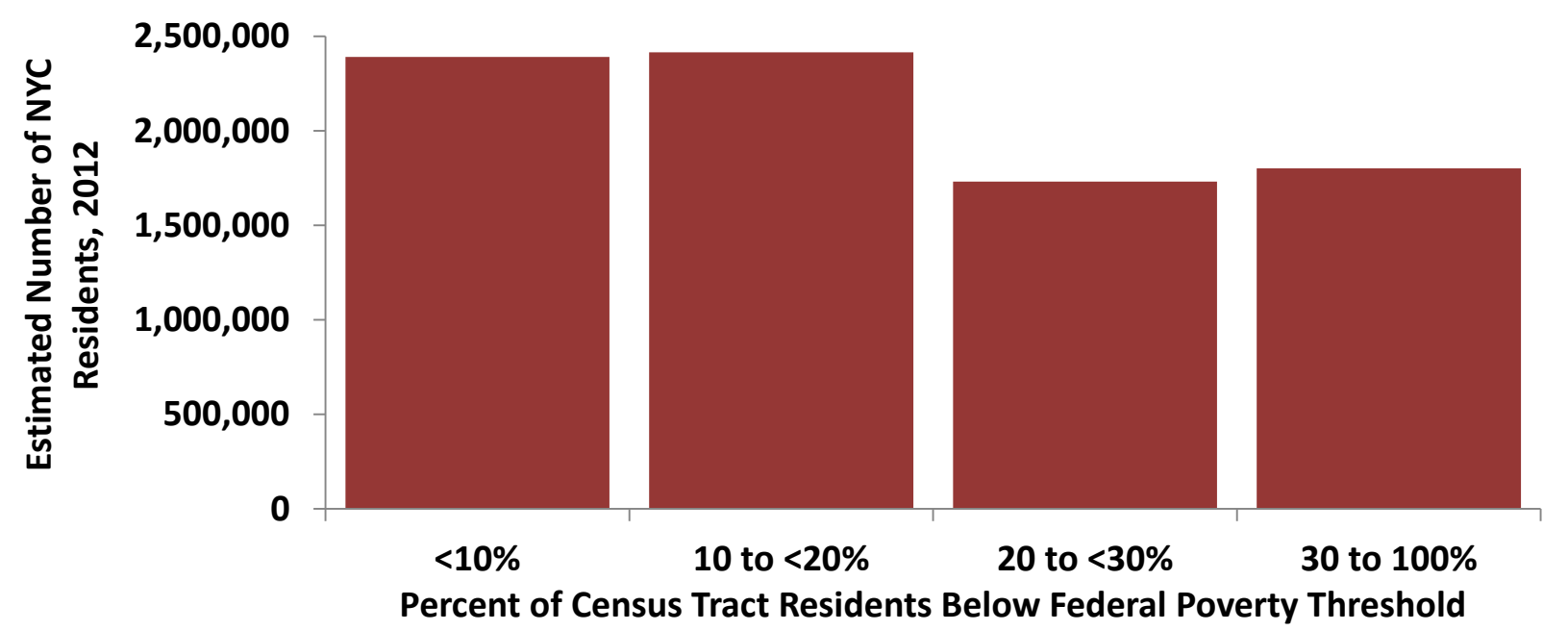
1. Geographic Systems Section, Information Technology Division, New York City Department of City Planning. Geosupport Desktop Edition™ copyrighted by the New York City Department of City Planning. Version 14.1, release 14A (http://www.nyc.gov/html/dcp/html/bytes/gdeguide.shtml). 2014.

Neighborhood Poverty

- Percent of census tract residents with incomes below federal poverty level, per American Community Survey (ACS)
  - For diagnoses 2006–2008: ACS 2006–2010
  - For diagnoses 2009: ACS 2007–2011
  - For diagnoses 2010–2013: ACS 2008–2012
- Four categories of percent below federal poverty level:
  - 0 to <10% (low poverty areas)
  - 10 to <20%
  - 20 to <30%
  - 30 to 100% (very high poverty areas)
    - Different from PHDGP cut-points: 0 to <5%, 5 to <10%, 10 to <20%, 20 to 100%

Overall Population Description

- Number of residents by poverty level



- Most common diseases, on average each year:
  - >9,000 newly reported cases each of chronic hepatitis B and C
  - >4,000 cases each of laboratory-confirmed influenza and respiratory syncytial virus
  - >1,000 cases each of Lyme disease, campylobacteriosis, and salmonellosis

Undomiciled Population

- May not appropriately geocode
- Identified undomiciled population
  - "Patient homeless at time of report?" field
  - Scanned address for keywords (undom, homeless)
  - Matched address at time of report against list of homeless shelters and address of a particular post office[1]
- Assigned to highest poverty category

[1] http://www.nydailynews.com/news/homeless-mail-general-delivery-article-1.276023

Incarcerated Population

- Incarcerated individuals need to be excluded :
  - Represent non-negligible proportion of cases for some diseases
  - Individuals living in group quarters do not contribute to area-based poverty measures
  - Home addresses when not incarcerated unavailable
- Identified incarcerated population
  - Scanned address for keywords (jail, prison)
  - Matched address at time of report against list of jails/prisons
  - Used Real Property Assets Database building classification code Y3 (Dept. of Finance)

Building Code	Description
Y1	FIRE DEPARTMENT
Y2	POLICE DEPARTMENT
Y3	PRISON, JAIL, HOUSE OF DETENTION
Y4	MILITARY AND NAVAL INSTALLATION
Y5	DEPARTMENT OF REAL ESTATE
Y6	DEPARTMENT OF SANITATION
Y7	DEPARTMENT OF PORTS AND TERMINALS
Y8	DEPARTMENT OF PUBLIC WORKS
Y9	DEPARTMENT OF ENVIRONMENTAL PROTECTION

Geocoding Results

- Across 53 diseases: 286,132 cases
  - 5,640 (2.0%) undomiciled → assigned to highest poverty level
  - 7,039 (2.5%) incarcerated → excluded from analysis
    - 91% of incarcerated cases were chronic hepatitis C
- Among 279,093 cases eligible for analysis, 252,395 (90.4%) successfully geocoded to a Census 2010 tract
- Non-geocodable cases more likely to be (chi-square $p<0.0001$):
  - Male
  - 25-64 years-old
  - Bronx residents
  - Cases of diseases not routinely investigated

Methods: Age-Adjusted Average Annual Rates

- Age may confound disease-poverty associations
  - Baseline disease risk and/or probability of diagnostic work-up (e.g., stool culture) may be associated with age
  - Poverty groups may have different underlying age structures
- Grouped cases into 7 age categories
  - <5, 5–14, 15–24, 25–44, 45–64, 65–74, ≥75 years
- Calculated age-adjusted rates by poverty level
  - Denominator: DOHMH intercensal population estimates
  - Direct standardization to US 2000 standard population[1]
  - 95% confidence intervals based on gamma distribution[2]

[1] http://www.cdc.gov/nchs/data/statnt/statnt20.pdf

[2] http://www.hsph.harvard.edu/thegeocodingproject/webpage/monograph/step_5.htm

Number 20 • January 2001 **Reprinted**

HEALTHY PEOPLE 2010 **Statistical Notes**

CDC SAFER·HEALTHIER·PEOPLE™

From the CENTERS FOR DISEASE CONTROL AND PREVENTION/National Center for Health Statistics

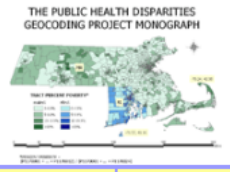
Age Adjustment Using the 2000 Projected U.S. Population

Richard J. Klein, M.P.H., and Charlotte A. Schoenborn, M.P.H.

http://www.cdc.gov/nchs/data/statnt/statnt20.pdf

Table 2: Selected age distributions and age-adjustment weights based on the 2000 projected U.S. population

Age	Population in thousands	Adjustment weight
Distribution #1		
All ages	**274,634**	**1.000000**
Under 1 year	3,795	0.013818
1–4 years	15,192	0.055317
5–14 years	39,977	0.145565
15–24 years	38,077	0.138646
25–34 years	37,233	0.135573
35–44 years	44,659	0.162613
45–54 years	37,030	0.134834
55–64 years	23,961	0.087247
65–74 years	18,136	0.066037
75–84 years	12,315	0.044842
85 years and over	4,259	0.015508
Distribution #2		
All ages	**274,634**	**1.000000**
Under 12 years	47,165	0.171738
12–19 years	31,619	0.115131
20–29 years	35,979	0.131007
30–39 years	41,691	0.151806
40–49 years	42,285	0.153968
50–59 years	30,531	0.111170
60–69 years	20,064	0.073057
70–79 years	16,141	0.058773
80 years and over	9,159	0.033350
Distribution #3		
All ages	**274,634**	**1.000000**
Under 18 years	70,783	0.257736
18–44 years	108,150	0.393797
45–54 years	37,030	0.134834
55–64 years	23,961	0.087247
65–74 years	18,136	0.066037
75 years and over	16,574	0.060349
Distribution #4		
All ages	**274,634**	**1.000000**
Under 18 years	70,783	0.257736
18–44 years	108,150	0.393797
45–64 years	60,991	0.222081
65–74 years	18,136	0.066037
75 years and over	16,574	0.060349
Distribution #5		
2 years and over	**267,080**	**1.000000**
2–5 years	15,329	0.057395
6–11 years	24,282	0.090917
12–19 years	31,619	0.118388
20–29 years	35,979	0.134712
30–39 years	41,691	0.156099
40–49 years	42,285	0.158323
50–59 years	30,531	0.114314
60–69 years	20,064	0.075124
70–79 years	16,141	0.060435
80 years and over	9,159	0.034293
Distribution #6		
2 years and over	**267,080**	**1.000000**
2–17 years	63,229	0.236742
18–44 years	108,150	0.404935
45–54 years	37,030	0.138647
55–64 years	23,961	0.089715
65–74 years	18,136	0.067905
75 years and over	16,574	0.062056



The Public Health Disparities Geocoding Project Monograph

Geocoding and Monitoring US Socioeconomic Inequalities in Health: An introduction to using area-based socioeconomic measures

STEP BY STEP COMPARISON

A step by step comparison of each task of the Case Example, the relevant section of Analytic Methods, and sample SAS code

(click here for a pdf version of all 8 steps)

Step 5:
For each category of CT poverty, calculate the age-standardized incidence rate, using the year 2000 standard million.

CASE EXAMPLE

In order to do this:

a. Aggregate the numerator and denominator within each age X CT poverty stratum, across all census tracts.

b. Exclude cases and denominator where CT poverty is missing.

c. Merge with the year 2000 standard million in five age categories.

d. Calculate the age-standardized incidence rate standardized to the year 2000 standard million, and the corresponding "gamma" confidence intervals for the direct standardized rates.

ANALYTIC METHODS

1. Age-standardized incidence rates

The standard practice of public health departments in reporting population rates of mortality and disease incidence is to calculate age-standardized rates, which facilitates comparisons between regions or subgroups of interest. The age-standardized rate is interpretable as the rate that would be observed in a population if that population had the same age distribution as a given reference population. Direct standardized rates are obtained by applying the age-specific incidence rates observed in the area or subgroup of interest to a standard age distribution, such as the year 2000 standard million.[1]

For our project, we used five broad age categories to age standardize, in order to obtain more stable rates in each age stratum, particularly for outcomes with sparse data.

If $cases_j$ represents the number of cases in age group j of the group or region of interest and pop_j represents the population associated with that age group, then the standardized rate IR_{st} for the group or region is

$$IR_{st} = \frac{\sum_j w_j \left(\frac{cases_j}{pop_j} \right)}{\sum_j w_j} = \frac{\sum_j w_j IR_j}{\sum_j w_j}$$

SAS PROGRAMMING

click here to download SAS program

```
**************************************************************
CREATE DATASET WITH STANDARD MILLION FOR AGE STANDARDIZATION
(IN FIVE CATEGORIES)
0-14
15-24
25-44
45-64
65+
**************************************************************;
data stdrd ;
input agecat y1940 y1970 y1980 y1990 y2000 ;

cards ;
1 250416 284926 226401 215383 214700
2 181677 174405 187542 147860 138646
3 301303 236183 276838 324695 298186
4 198105 205746 196440 186446 222081
5 68499 98740 112779 125616 126387
;

RUN ;


PROC SORT DATA=Step4 ;
BY AGECAT CINDPOV ;
run ;
```

http://www.hsph.harvard.edu/thegeocodingproject/webpage/monograph/step_5.htm

2. Confidence intervals for directly standardized rates

Traditional confidence limits for the direct standardized rates are based on the normal distribution and require large cell counts. In our analyses, we found that they can also occasionally result in "impossible" lower limits that are less than zero. Because of this, we adopted an alternate method for calculating the confidence limits based on the inverse gamma function.[2] This method assumes that the direct standardized rate is a linear combination of independent Poisson random variables. Assuming that this linear combination is also follows a Poisson distribution, the age-standardized rate $E(X) = x$ follows a gamma distribution $\Gamma(a,b)$ as follows:

$$X \sim \Gamma\left(\frac{x^2}{v}, \frac{v}{x}\right)$$

```
**********
LOWER 95% GAMMA INTERVAL

LGAM gives the 95% gamma interval using the formula given by Fay and
Feuer.
LGAM2 gives the 95% gamma interval using the formula given Anderson and
Rosenberg.

***NOTE: FOR LGAM2 AND UGAM2, HAVE NOW PROGRAMMED OPTIONS
FOR IRW=0 VARPYW=0
I.E. USE INVERSE CHI SQUARE DISTRIBUTION CINV(0.975,2) AND DIVIDE
BY DENOMINATOR
TO GET UPPER LIMIT ON RATE

References:
Fay MP, Feuer EJ. Confidence intervals for directly standardized rates:
a method based on the gamma distribution. Statistics in Medicine 1997,16:791-
801.

Anderson RN, Rosenberg HM. Age standardization of death rates:
implementation of the year 2000 standard.
National Vital Statistics Reports: Vol 37, No. 3. Hyattsville, MD:
National Center for Health Statistics, 1998.
**********;

LGAM=(VARPYW/(2*IRW)) * CINV(0.025,((2*(IRW**2))/VARPYW)) ;

IF IRW=0 AND VARPYW=0 THEN DO ;
LGAM2=0 ;
END ;
ELSE LGAM2=(VARPYW/IRW) * GAMINV(0.025,((IRW**2)/VARPYW)) ;

**********
UPPER 95% GAMMA

UGAM gives the 95% gamma interval using the formula given by Fay and
Feuer.
UGAM2 gives the 95% gamma interval using the formula given Anderson and
Rosenberg.

**********;

UGAM=((VARPYW + (WMAX**2))/2*(IRW+WMAX)) * CINV(0.975,((2*((IRW +
WMAX)**2))/(VARPYW + (WMAX**2)))) ;

IF IRW=0 AND VARPYW=0 THEN DO ;
UGAM2=(0.5 * CINV(0.975,2))/CRDEN ;
END ;
ELSE UGAM2=(VARPYW/IRW) * GAMINV(0.975,(((IRW**2)/VARPYW) + 1)) ;

**********
```

Methods: Disparity Measures

- Age-adjusted incidence rate ratios (IRRs) between highest vs. lowest poverty groups
- Chi-square (Cochran-Armitage) test for trend
  - Indicates step-wise increase or decrease in disease rates across poverty levels
  - Results not presented for diseases where data too sparse (expected disease counts <5 for ≥2 poverty levels)
- Population attributable fraction (PAF)
  - Fraction of all cases that might not have occurred without exposure to higher poverty neighborhoods[1]

[1] http://www.hsph.harvard.edu/thegeocodingproject/webpage/monograph/step_8.htm

6. Population Attributable Fraction

The population attributable fraction (PAF) is a useful summary measure for characterizing the public health impact of an exposure on population patterns of health and disease. It is defined as "the fraction of all cases (exposed and unexposed) that would not have occurred if exposure had not occurred."[8] For a polytymous exposure, the population attributable fraction is a weighted sum of the attributable fractions for each level of the exposure, with the weights defined by the case fractions (number of exposed cases divided by overall number of cases):

$$PAF = CF_1 \times \frac{RR_1 - 1}{RR_1} + CF_2 \times \frac{RR_2 - 1}{RR_2} + \ldots + CF_j \times \frac{RR_j - 1}{RR_j}$$

```
************************************************************
        STEP 8

        Using the age-stratified numerators and denominators from Step 4,
        calculate the age-stratum specific population attributable
        risk fractions and aggregate population attributable risk fraction
        over all age strata following the method of Hanley.

        Reference:
        JA Hanley, A heuristic approach to the formulas for population attributable
        fraction. J Epidemiol Community Health 2001,55:508-514.

        SOME NOTATION:

        Assume that the dataset provided has stratum specific numbers of
        cases (NUMER) and denominator (DENOM).

        Subscript i as age, j as covariate (in this case, CINDPOV)


        1. Sort data by AGECAT and CINDPOV.

        2. Calculate the quantities
                NUMERi+ = SUMj(NUMERij)
                RATEiREF  = rate in the reference group
           and save them in a dataset.

        3. Merge the quantities from (2) with the dataset and calculate

                (a) rate                RATEij=NUMERij/DENOMij
                (b) case fractions      CFij = NUMERij/NUMERi+
                (c) rate ratio          RRij = RATEij/RATEiREF

        4. Calculate the (age)stratum-specific population attributable risk fraction

           AFPi = CFi1 * (RRi1-1)/RRi1 + CFi2 * (RRi2-1)/RRi2 + ... CFij * (RRij-1)/RRij

        5. Calculate the grand total of cases NUMER++ to use to calculate age-specific weights.

        6. Determine age-stratum specific weights from the case distribution:

                w_i = NUMERi+/NUMER++

           and calculate the aggregated AFPagg = SUM(w_i * AFPi)

************************************************************;

*********************************
STEP 8a SUM OVER AREAKEY INTO STRATA BY AGECAT AND CINDPOV
*********************************;
```

Example: PAF for *Salmonella* for one Age Stratum

Age	Poverty level	SAL count	Proportion of cases in poverty level (case fraction)	Person-years	Rate	Rate ratio	Attributable fraction
<5	<10% (ref)	692	0.19	1,077,452	0.00064	1	0
	10–<20%	934	0.26	1,137,128	0.00082	1.28	0.0569
	20–<30%	896	0.25	907,811	0.00099	1.54	0.0875
	30–100%	1,055	0.29	1,106,880	0.00095	1.48	0.0962
Total		3,577	1				PAF=0.2406

$$PAF = CF_1 \times \frac{RR_1 - 1}{RR_1} + CF_2 \times \frac{RR_2 - 1}{RR_2} + \ldots + CF_j \times \frac{RR_j - 1}{RR_j}$$

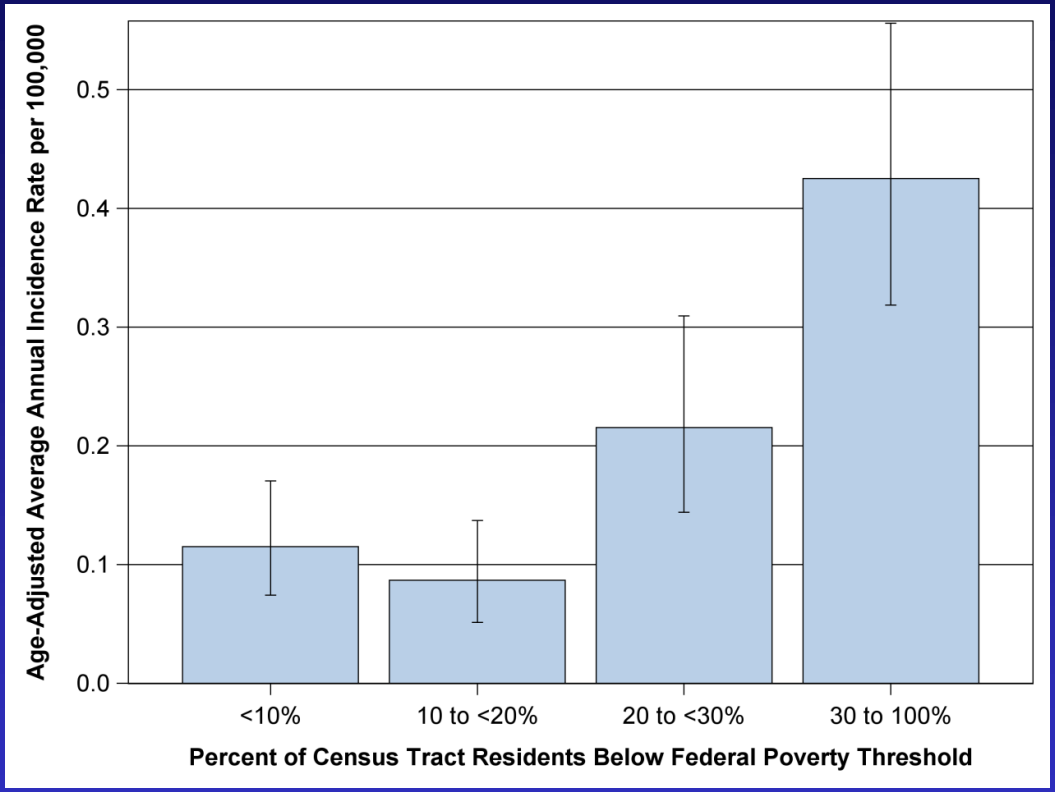
PAF Example: *Salmonella*

Age Group	SAL count	Proportion of cases in age stratum (case fraction)	PAF	Weighted PAF
<5	3,577	0.39	0.241	0.0931
5–14	1,352	0.15	0.005	0.0007
15–24	804	0.09	-0.214	-0.0186
25–44	1,580	0.17	-0.022	-0.0038
45–64	1,143	0.12	0.220	0.0271
65–74	380	0.04	0.124	0.0051
≥75	426	0.05	0.165	0.0076
Total	9,262	1		PAFagg=0.111

Diseases Associated with High Poverty (n=18)

Disease	Number of Cases	Years in Study Period	Percent Geocoded	IRR (95% CI) for Highest vs. Lowest Poverty	P-value (Test for Trend)	PAF
Rickettsialpox	134	8	94	3.69 (2.29, 5.95)	<.0001	0.39
Hepatitis C, chronic	75,929	8	87	3.58 (3.50, 3.66)	<.0001	0.45
Malaria	1,695	8	93	3.48 (2.97, 4.08)	<.0001	0.52
Hepatitis B, chronic	74,664	8	89	3.28 (3.20, 3.36)	<.0001	0.52
Streptococcus pneumoniae	5,717	7	95	2.61 (2.42, 2.81)	<.0001	0.35
Vancomycin-intermediate *Staphylococcus aureus*	56	6	88	2.51 (1.18, 5.35)	0.02	0.30
Group A *Streptococcus*	1,640	8	96	2.33 (2.03, 2.68)	<.0001	0.30
Shigellosis	3,387	8	91	2.31 (2.08, 2.58)	<.0001	0.40
Group B *Streptococcus*	343	8	97	2.29 (1.65, 3.18)	<.0001	0.41
Legionellosis	1,599	8	98	2.04 (1.79, 2.34)	<.0001	0.24
Neisseria meningitidis	218	8	96	2.02 (1.38, 2.97)	<.0001	0.27
Hepatitis B, acute	707	8	95	1.94 (1.56, 2.41)	<.0001	0.29
Haemophilus influenzae	877	8	91	1.81 (1.50, 2.18)	<.0001	0.17
RSV	26,479	6	93	1.78 (1.71, 1.84)	<.0001	0.28
Cryptosporidiosis	842	8	99	1.67 (1.38, 2.03)	<.0001	0.22
Dengue	684	8	94	1.54 (1.24, 1.92)	<.0001	0.17
Salmonellosis	9,802	8	95	1.27 (1.20, 1.35)	<.0001	0.11
Influenza	38,776	8	92	1.17 (1.14, 1.21)	<.0001	0.05

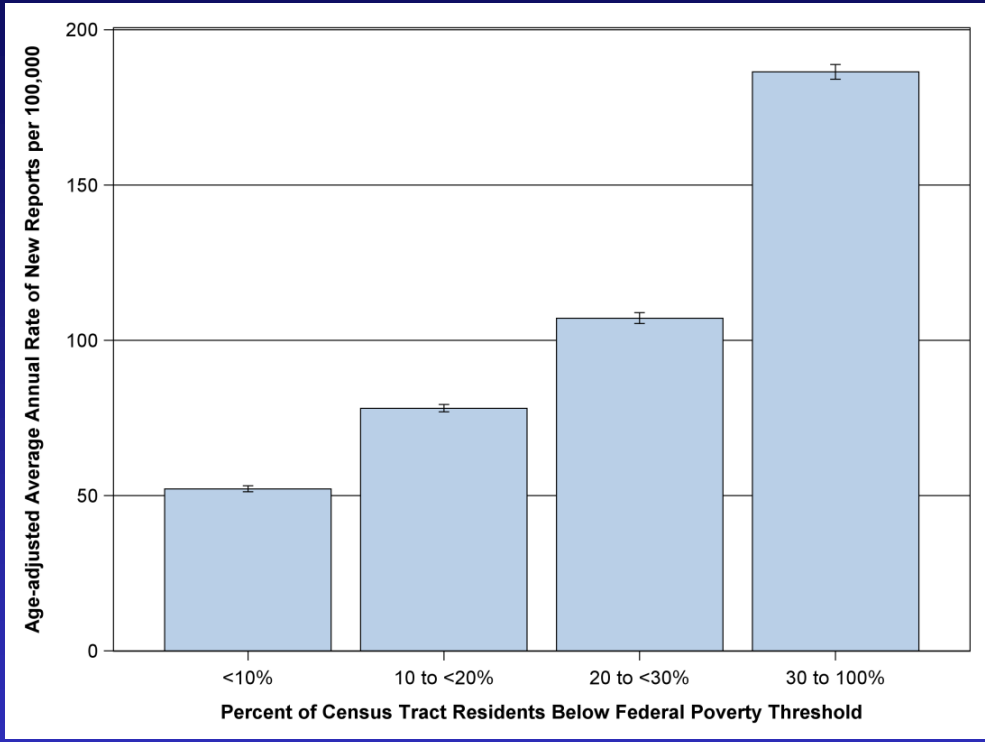
Rickettsialpox



- Transmitted to humans from mice via mites
- Residents of high poverty areas likely have greater exposure to house mice[1]

[1] Paddock CD, Zaki SR, Koss T, et al. Rickettsialpox in New York City: a persistent urban zoonosis. *Annals of the New York Academy of Sciences.* Jun 2003;990:36-44.

Chronic Hepatitis C



- Association driven by older age groups (particularly those born 1945–1965) and by injection drug use[1,2]
- Prevalence is very high[3]
- Prevention programs need strengthening[4]
- Curative treatment available

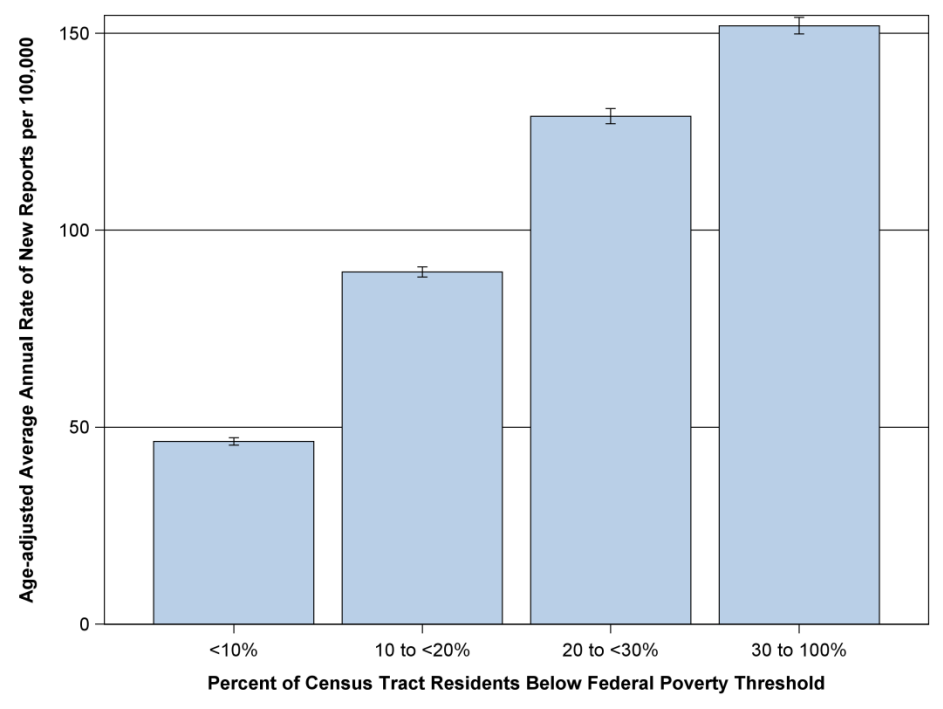
[1] Prussing C, et al. Hepatitis C surveillance among youth and young adults in New York City, 2009–2013. *Journal of urban health.* 2014.

[2] Baumgartner J, et al. Hepatitis B and C surveillance report: New York City, 2013. Available from http://www.nyc.gov/html/doh/html/data/cd-hepabc-reports.shtml. 2015.

[3] Balter S, et al. Estimating the prevalence of hepatitis C infection in New York City using surveillance data. *Epidemiology and infection.* 2014.

[4] Hagan H. Agent, host, and environment: hepatitis C virus in people who inject drugs. *Journal of infectious diseases.* 2011.

Chronic Hepatitis B

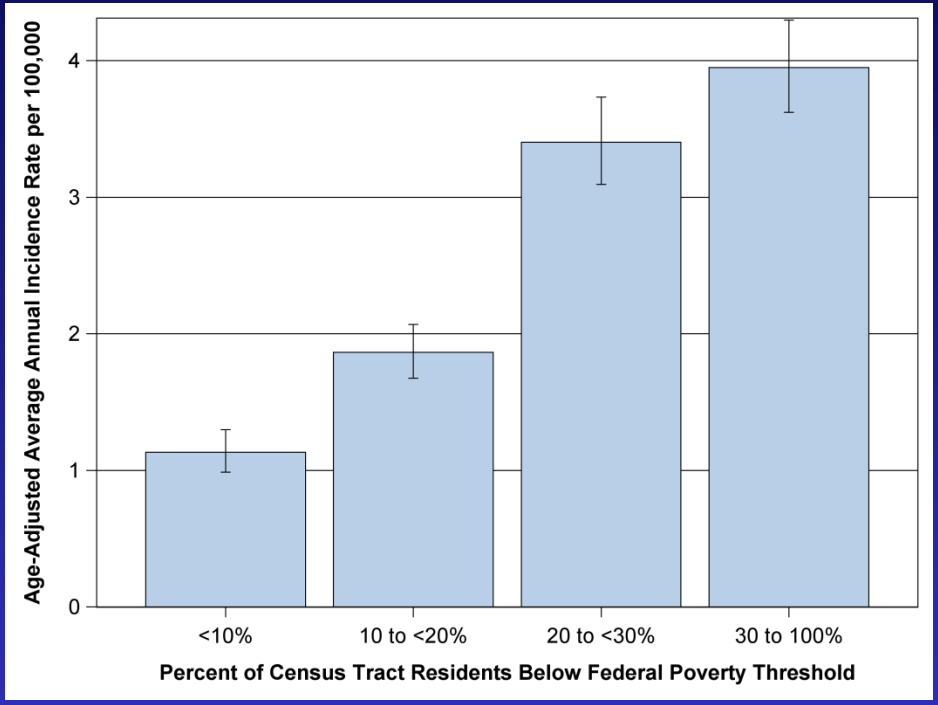


- **Largely driven by immigration patterns from China[1,2]**

[1] McGibbon E, et al. Surveillance for chronic hepatitis B virus infection - New York City, June 2008-November 2009. *MMWR.* 2012.

[2] France AM, et al. Estimating the prevalence of chronic hepatitis B virus infection--New York City, 2008. *Journal of urban health.* 2012.

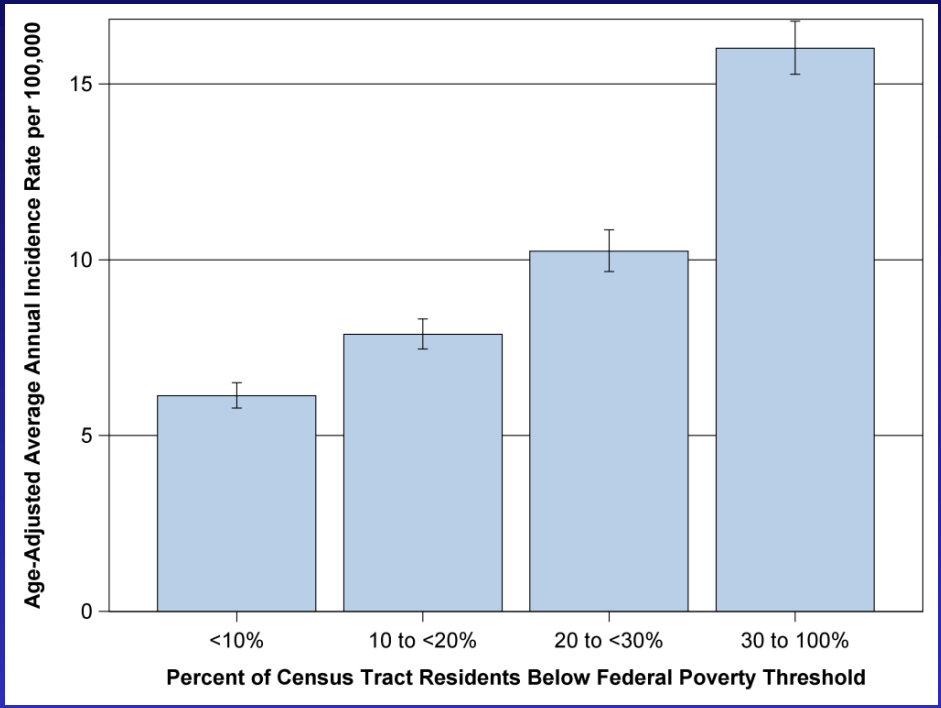
Malaria



- All cases were travel-related
  - Majority infected while traveling to visit friends and relatives or before recent immigration to NYC[1]
- Association with high poverty neighborhoods likely reflects immigration patterns
- Antimalarial prophylaxis should be promoted among at-risk travelers in these neighborhoods

[1] Adamson R, et al. Epidemiology and burden of hepatitis A, malaria, and typhoid in New York City associated with travel: implications for public health policy. *American journal of public health.* 2010.

Invasive Pneumococcal Disease

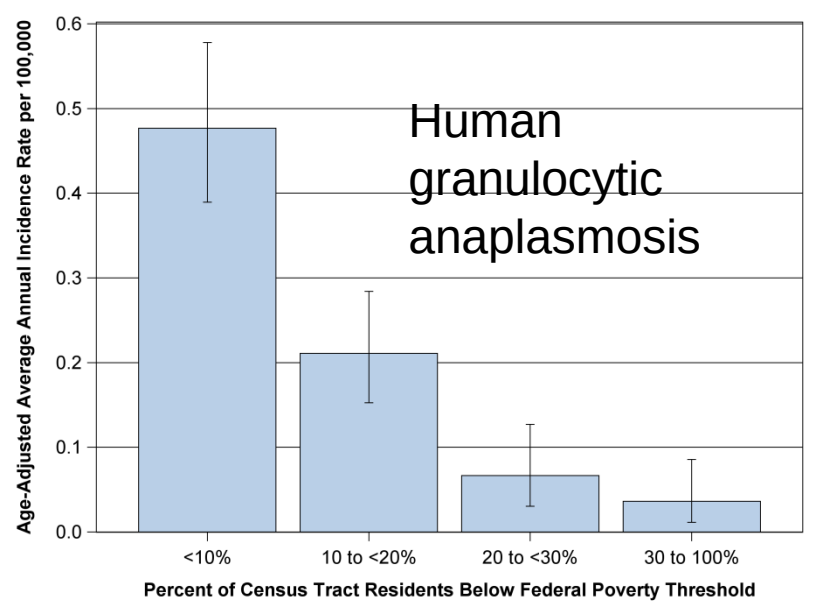


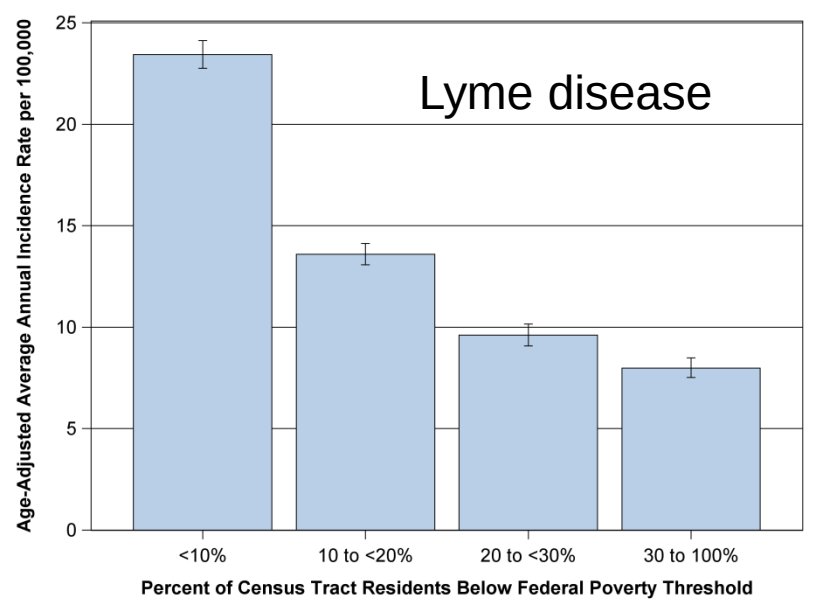
- Association with poverty previously reported
  - New York City[1]
  - Connecticut[2]
  - Areas across nine states (CDC's ABC surveillance/EIP network)[3]
- Might be related to lower pneumococcal vaccination rates, higher rates of underlying medical conditions and smoking, and/or crowding in higher poverty areas

[1] Dentinger C, et al. Invasive Pneumococcal Disease Surveillance in New York City. *Epi Data Brief.* 2011.
[2] Soto K, et al. Changing disparities in invasive pneumococcal disease by socioeconomic status and race/ ethnicity in Connecticut, 1998-2008. *Public Health Rep.* 2011.
[3] Burton DC, et al. Socioeconomic and racial/ethnic disparities in the incidence of bacteremic pneumonia among US adults. *American journal of public health.* 2010.

Diseases Associated with Low Poverty (n=11)

Disease	Number of Cases	Years in Study Period	Percent Geocoded	IRR (95% CI) for Highest vs. Lowest Poverty	P-value (Test for Trend)	PAF
Anaplasmosis, human granulocytic	172	8	95	0.08 (0.03, 0.19)	<.0001	-0.81
Ehrlichiosis, human monocytic	82	8	99	0.10 (0.04, 0.29)	<.0001	-0.66
Hemolytic uremic syndrome	33	8	100	0.19 (0.06, 0.68)	0.004	-0.34
Babesiosis	340	8	95	0.20 (0.13, 0.32)	<.0001	-0.68
West Nile neuroinvasive disease	109	8	95	0.22 (0.09, 0.51)	<.0001	-0.33
Transmissible spongiform encephalopathies	51	8	80	0.24 (0.07, 0.80)	0.001	-0.49
Vibrio species (non-cholera)	132	8	96	0.27 (0.14, 0.52)	<.0001	-0.49
Cyclosporiasis	140	8	100	0.31 (0.17, 0.55)	<.0001	-0.51
Lyme disease	10,763	8	91	0.34 (0.32, 0.36)	<.0001	-0.59
Shiga toxin-producing *Escherichia coli*	560	8	96	0.52 (0.41, 0.68)	<.0001	-0.30
Giardiasis	6,958	8	98	0.88 (0.82, 0.94)	0.0001	-0.05





Tickborne Diseases

- Includes human granulocytic anaplasmosis, Lyme disease, babesiosis, human monocytic ehrlichiosis
- Likely reflects a population wealthy enough to travel to areas outside of NYC where infected vectors are prevalent[1]

[1] Abdool AJ, et al. Lyme disease in New York City – Is it locally acquired? Poster presented at 6th International Conference on Emerging Infectious Diseases, March 16–19, 2008; Atlanta, GA.

Remaining Diseases

Null association (n=12):

- Amebiasis
- Campylobacteriosis
- Hepatitis A
- Hepatitis C, acute
- Leprosy
- Listeriosis
- Norovirus, laboratory-confirmed
- Rotavirus, laboratory-confirmed
- Rocky Mountain spotted fever
- Paratyphoid fever
- Typhoid fever
- Yersiniosis

Data too sparse (n=12):

- *Anthrax*
- *Botulism, infant*
- *Botulism, foodborne or other*
- *Brucellosis*
- *Cholera*
- *Ehrlichiosis, not otherwise specified*
- *Lymphocytic choriomeningitis virus*
- *Leptospirosis*
- *Q fever*
- *Trichinosis*
- *Toxic shock syndrome*
- *Tularemia*

Primary Limitation: Diagnostic Bias

- Residence in high poverty areas may affect probability of seeking care and receiving a diagnosis
  - Could be less likely
    - Inadequate health insurance
    - Expensive or specialized diagnostic testing
  - Could be more likely
    - Disproportionately use emergency departments, where testing more common than primary care settings
    - Greater access to disease screening

Additional Limitations

- Patients who successfully geocode may be non-representative
- Data sparse for some diseases, poverty levels, and age groups
- Temporal relationship can be unclear for chronic infections
- Multiple testing

Summary

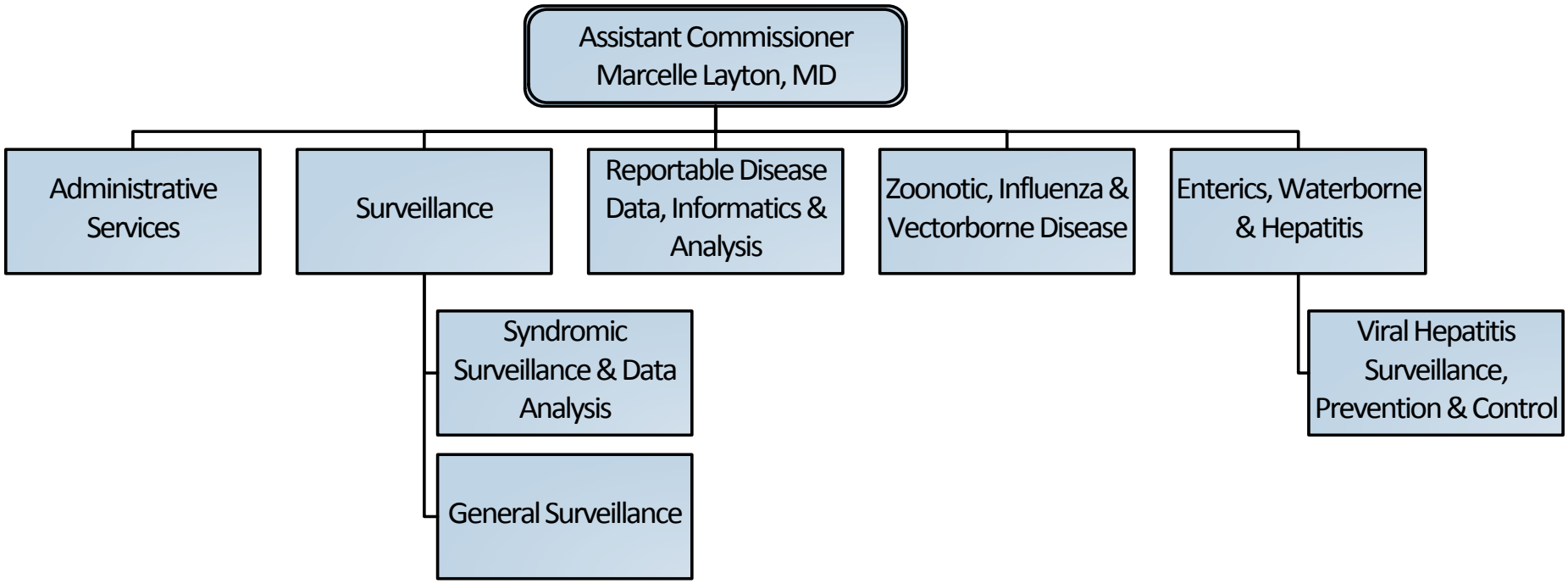
- Of 41 diseases with adequate sample size:
  - 18 diseases (44%) associated with high poverty areas
  - 11 diseases (27%) associated with low poverty areas
- Findings may be useful for:
  - Targeting outreach or educational efforts
  - Serving as a baseline for monitoring disparities over time and across jurisdictions
- Resources and prevention efforts should be targeted to high poverty areas

Next Steps

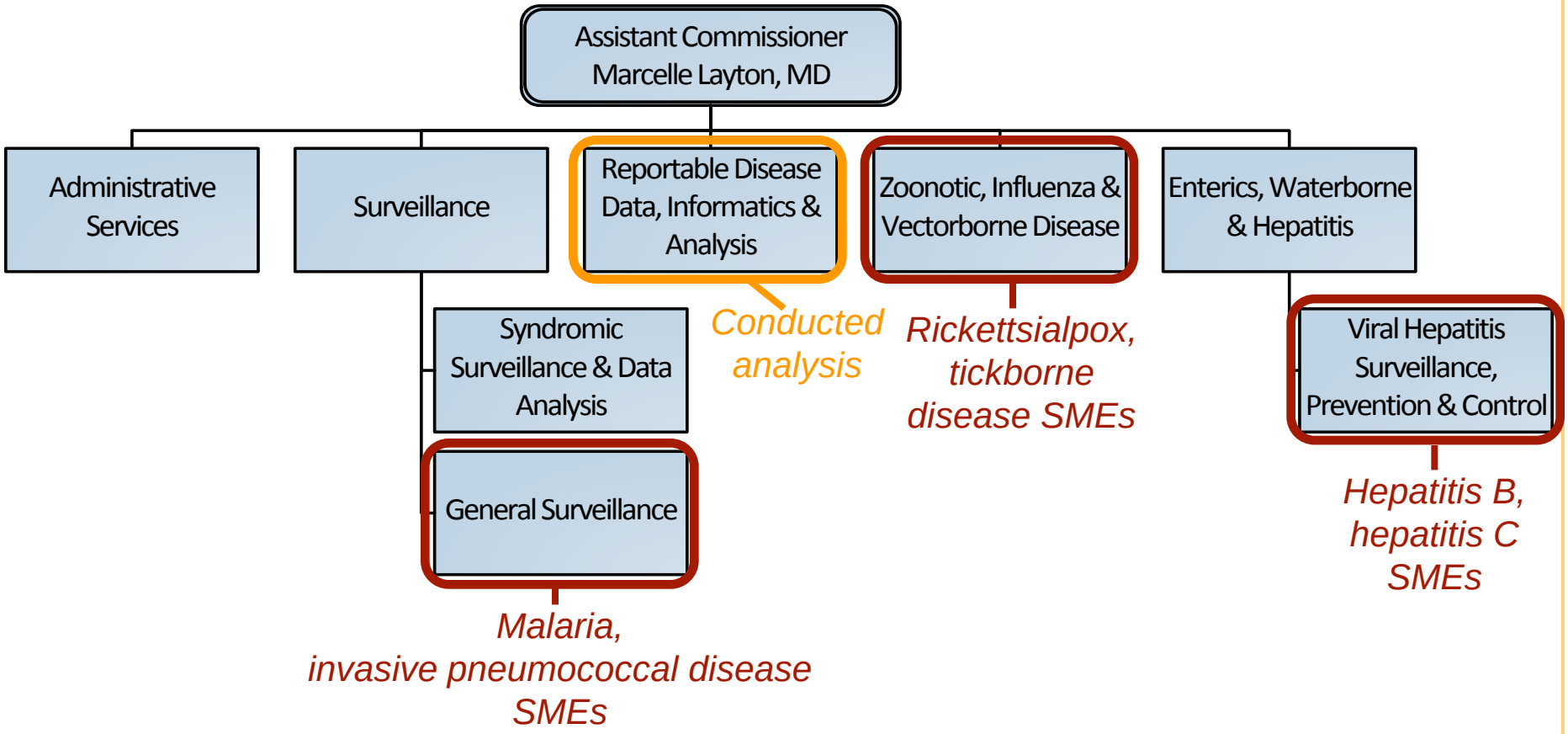
- Clarify subgroups at highest risk
  - Pathogen subtype
  - Patient age group, sex, race/ethnicity, household income, travel history, underlying medical conditions, environmental exposures
  - Adjust for household crowding
- Identify reasons for the observed associations
- Use findings to support programs to minimize inequities

Soliciting input for interpreting findings from subject matter experts

Bureau of Communicable Disease Organizational Chart



Bureau of Communicable Disease Organizational Chart



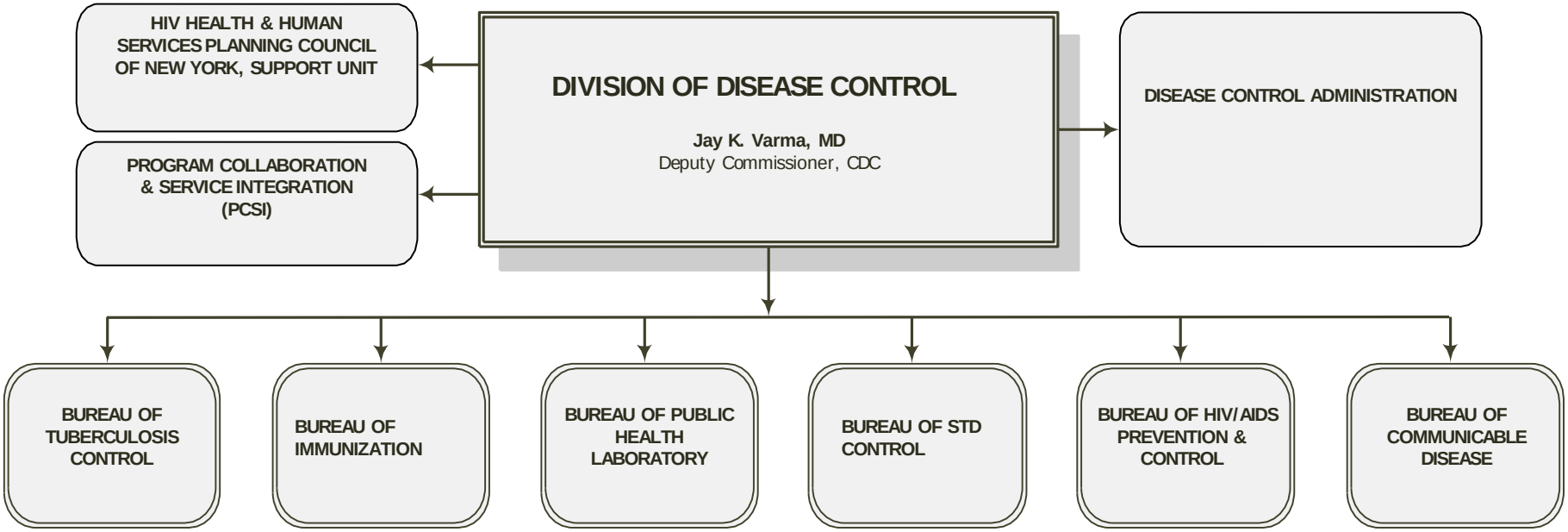
Soliciting Stakeholder Input

- Fora for presenting and discussing analyses in-progress
  - Bureau-wide: “Meeting of the Minds”
  - Agency-wide: “Epi Grand Rounds”
- Manuscript preparation
  - Circulated draft manuscript to disease reviewers
  - Cited prior relevant work by SMEs, including reports and posters
  - E-mails, side conversations
  - Agency clearance
  - Met with key SMEs prior to submission to clarify conclusions
- Challenges
  - Competing priorities
  - Timing

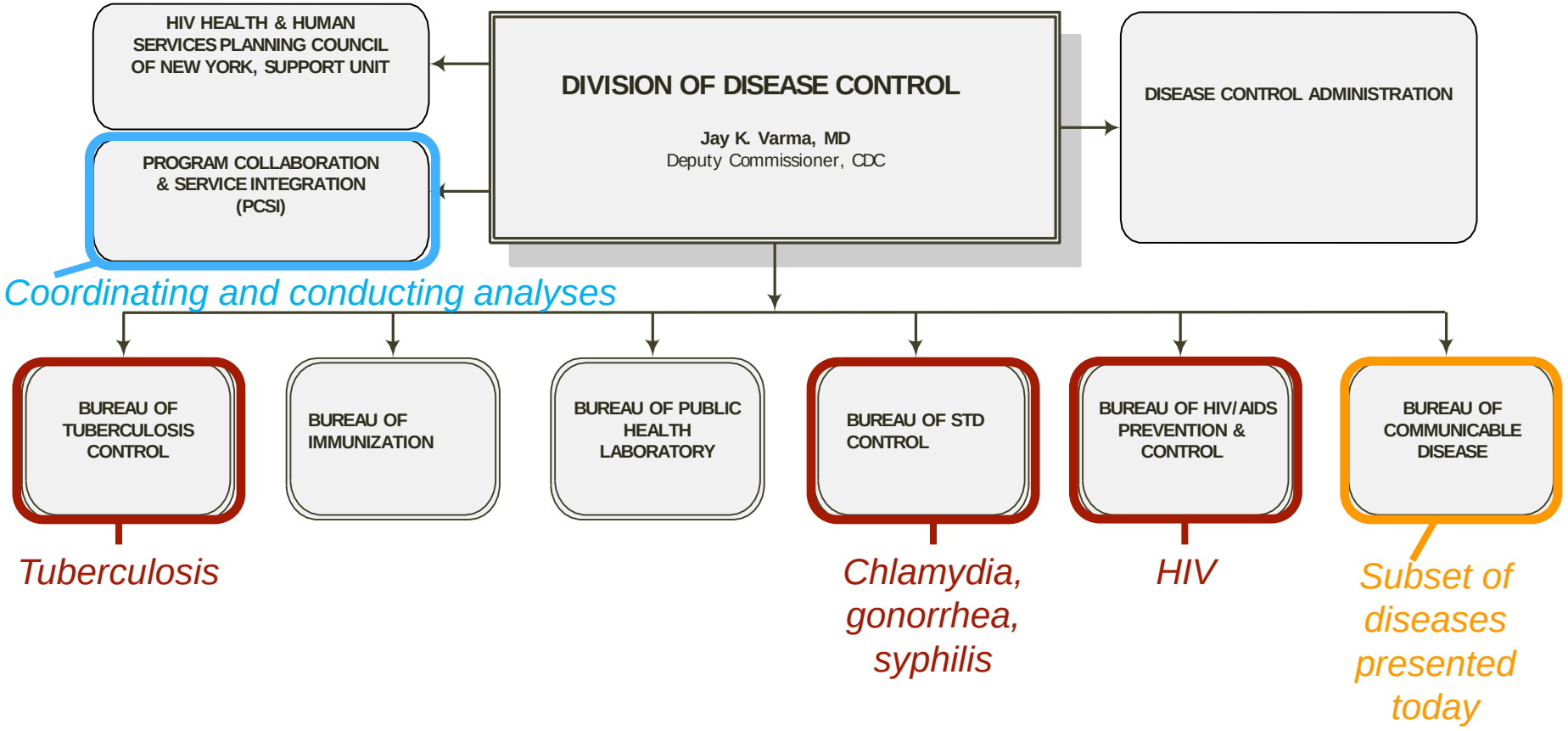
Expanding Beyond One Bureau

- Expanding analysis to also include HIV, tuberculosis, and selected sexually transmitted infections
- Epi Data Brief
  - Short (2-3 pages) publication
  - Highlights data findings from varying Health Department programs and projects

http://www.nyc.gov/html/doh/html/data/epidata.shtml



October 2014



Will solicit stakeholder input through PCSI Data Advisory Committee and Epi Data Brief review process.

October 2014

Selected Resources

- ✓ Public Health Disparities Geocoding Project
  - http://www.hsph.harvard.edu/thegeocodingproject/
- ✓ Toprani A, Hadler JL. Selecting and applying a standard area-based socioeconomic status measure for public health data: Analysis for New York City. NYC DOHMH: Epi Research Report, May 2013
  - http://www.nyc.gov/html/doh/downloads/pdf/epi/epiresearch-SES-measure.pdf
- ✓ Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Vol 20. Hyattsville, MD: National Center for Health Statistics; 2001
  - http://www.cdc.gov/nchs/data/statnt/statnt20.pdf
- ✓ Greene SK, Levin-Rector A, Hadler JL, Fine AD. Disparities in reportable communicable disease incidence by census tract-level poverty, New York City, 2006–2013 (Am J Public Health, provisionally accepted)
  - Corresponding author: sgreene4@health.nyc.gov


ANY
QUESTIONS
?

Extra Slides

Methods: Tests for Trend

- Set up 2x4 tables by multiplying age-adjusted rate by total population to derive age-adjusted counts of disease events and unaffected persons for each poverty level

Percent of census tract residents below federal poverty threshold	Age-adjusted counts of disease events	Age-adjusted counts of unaffected persons
<10%		
10 to <20%		
20 to <30%		
≥30%		

- Results not presented for diseases where data too sparse
  - Expected disease counts <5 for ≥2 poverty levels
  - SAS: "WARNING: 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test"
- Chi-square (Cochran-Armitage) test for trend
  - Indicates step-wise increase or decrease in disease rates across poverty levels

Diseases Not Associated with Poverty (n=12)

Disease	Number of Cases	Number of Years in Study Period	Percent Geocoded	IRR (95% CI) for Highest vs. Lowest Poverty Level	P-value (Test for Trend)	PAF
Amebiasis	3,584	8	96	0.96 (0.87, 1.07)	0.25	0.03
Campylobacteriosis	9,706	8	97	0.94 (0.88, 1.00)	0.61	0.06
Hepatitis A	771	8	95	0.86 (0.68, 1.08)	0.07	0.07
Hepatitis C, acute	69	8	93	1.01 (0.46, 2.21)	0.57	0.23
Leprosy	50	8	88	1.84 (0.74, 4.55)	0.47	0.48
Listeriosis	289	8	95	1.12 (0.80, 1.58)	0.55	0.06
Norovirus, laboratory-confirmed	233	6	87	1.26 (0.85, 1.88)	0.24	0.07
Rotavirus, laboratory-confirmed	673	6	94	1.11 (0.88, 1.40)	0.11	0.16
Rocky Mountain spotted fever	100	8	96	0.66 (0.38, 1.16)	0.09	-0.26
Paratyphoid fever	66	5	96	1.10 (0.48, 2.53)	0.87	0.32
Typhoid fever	394	8	97	1.31 (0.94, 1.83)	0.19	0.35
Yersiniosis	161	8	88	0.79 (0.48, 1.29)	0.23	-0.01