Big data, social determinants of health, and health inequities

Sandro Galea

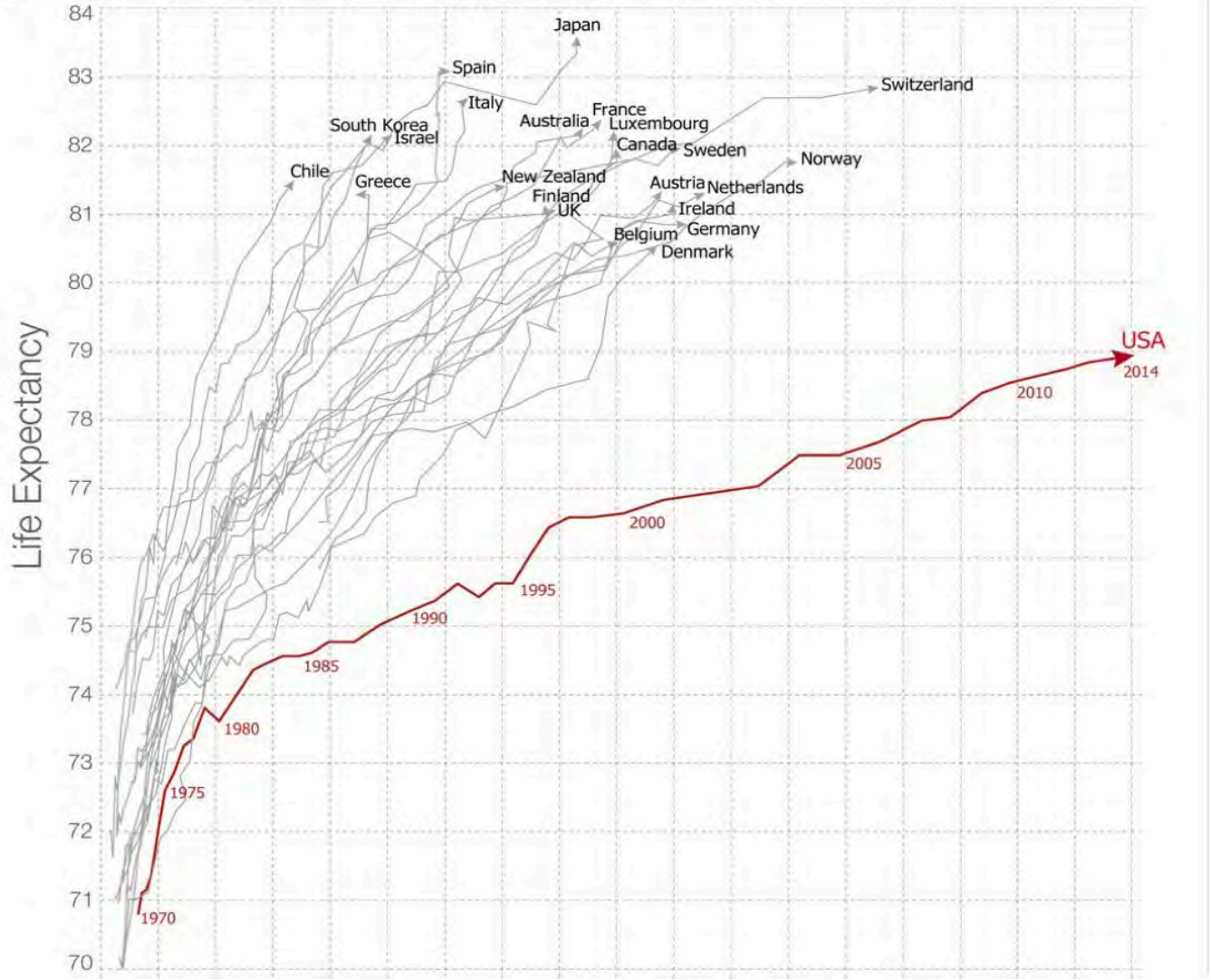
Boston University School of Public Health

1. The trouble with the population's health

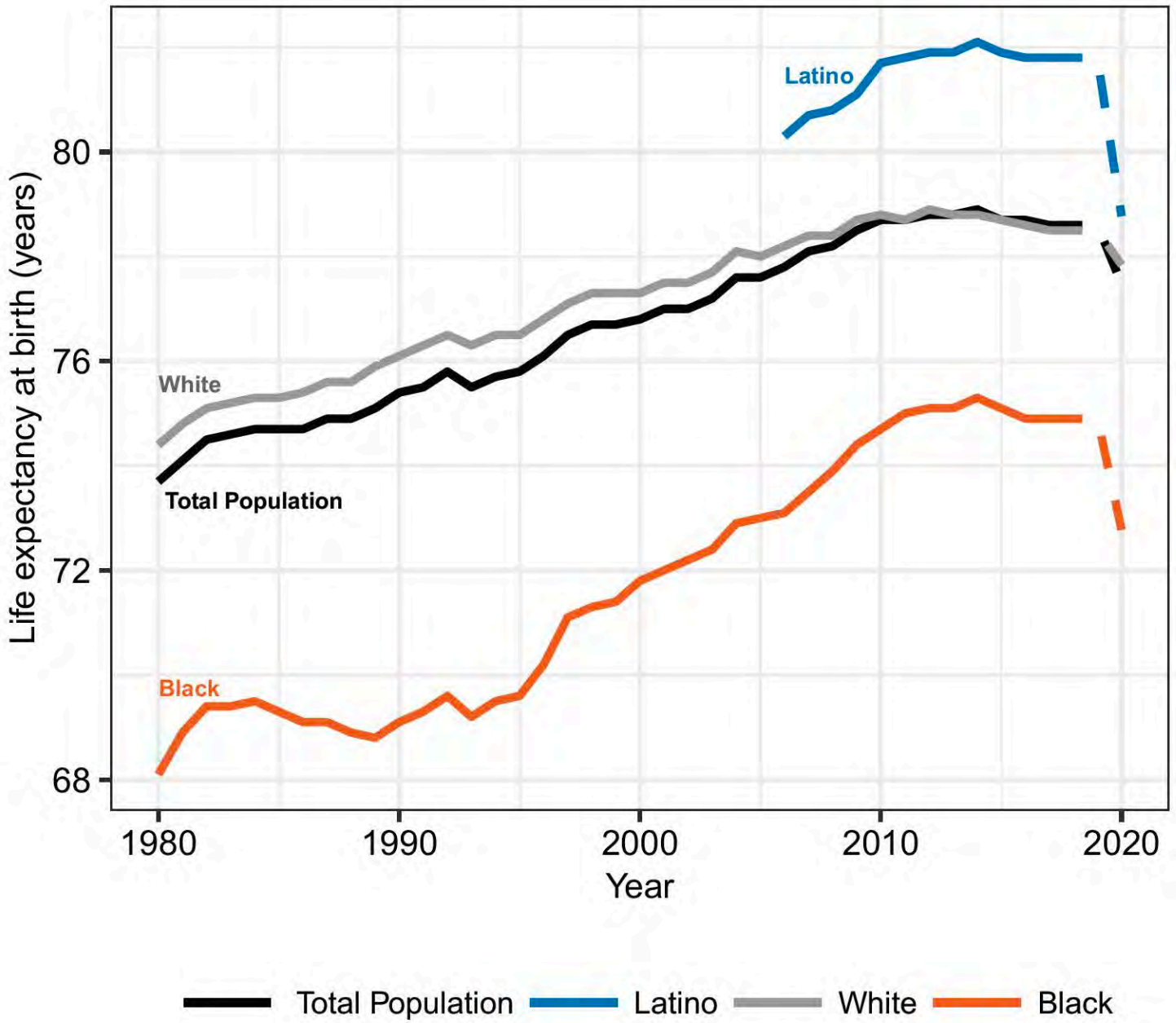
Life expectancy vs. health expenditure over time (1970-2014)

Our World in Data

Health spending measures the consumption of health care goods and services, including personal health care (curative care, rehabilitative care, long-term care, ancillary services and medical goods) and collective services (prevention and public health services as well as health administration), but excluding spending on investments. Shown is total health expenditure (financed by public and private sources).



Our World in Data. "The link between health spending and life expectancy: The US is an outlier."
<https://ourworldindata.org/the-link-between-life-expectancy-and-health-spending-us-focus> Accessed September 7, 2016



Andrasfay T, Goldman N. Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations. PNAS 2021 Vol. 118 No. 5 e2014746118

2. Understanding population health

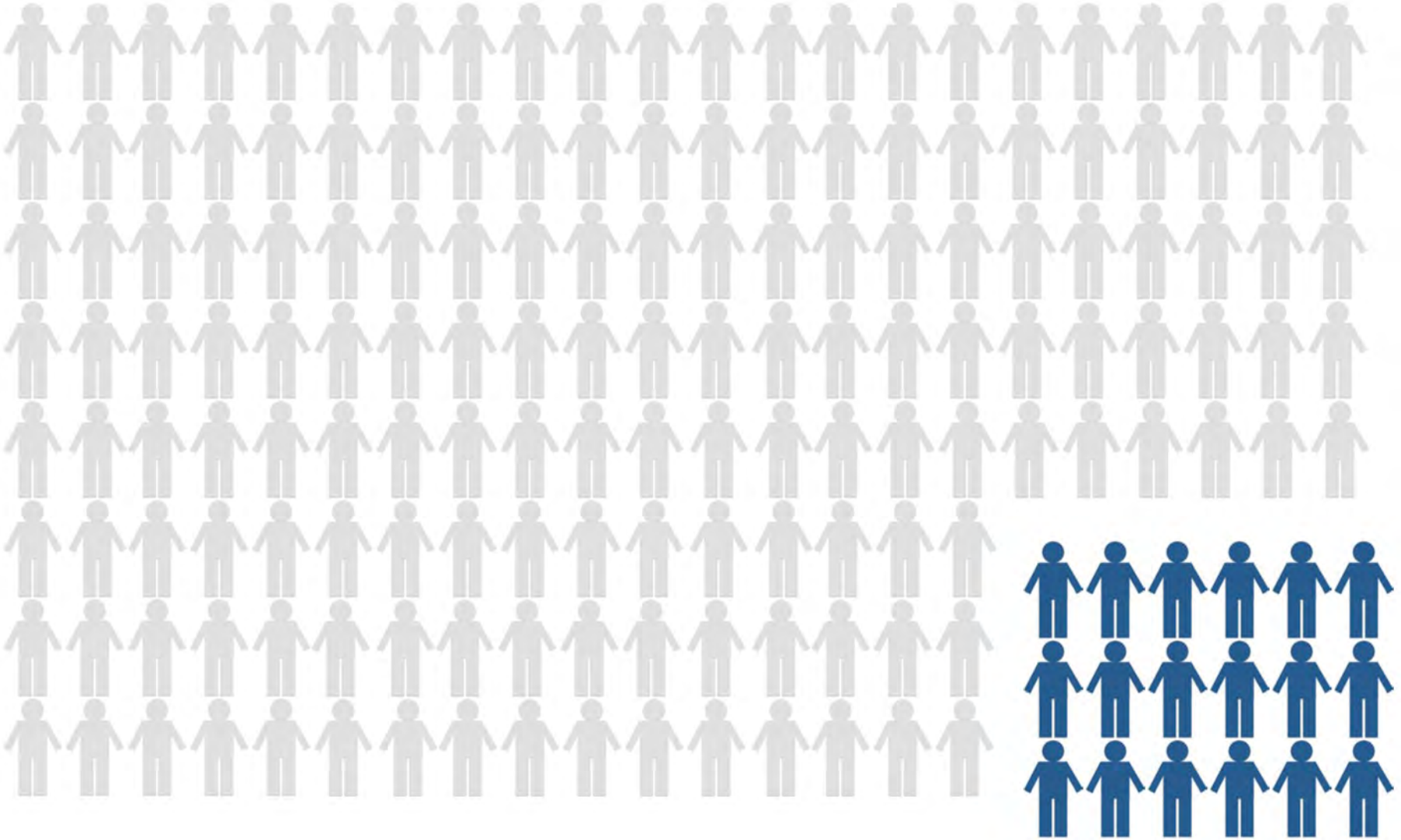
“The health outcomes of a group of individuals, including the distribution of such outcomes within the group”

Kindig D, Stoddart G. What is population health? American Journal of Public Health 2003; 93;380-383.

“The health outcomes of a group of individuals, including the distribution of such outcomes within the group”

Why? So that we may intervene

Kindig D, Stoddart G. What is population health? American Journal of Public Health 2003; 93;380-383.

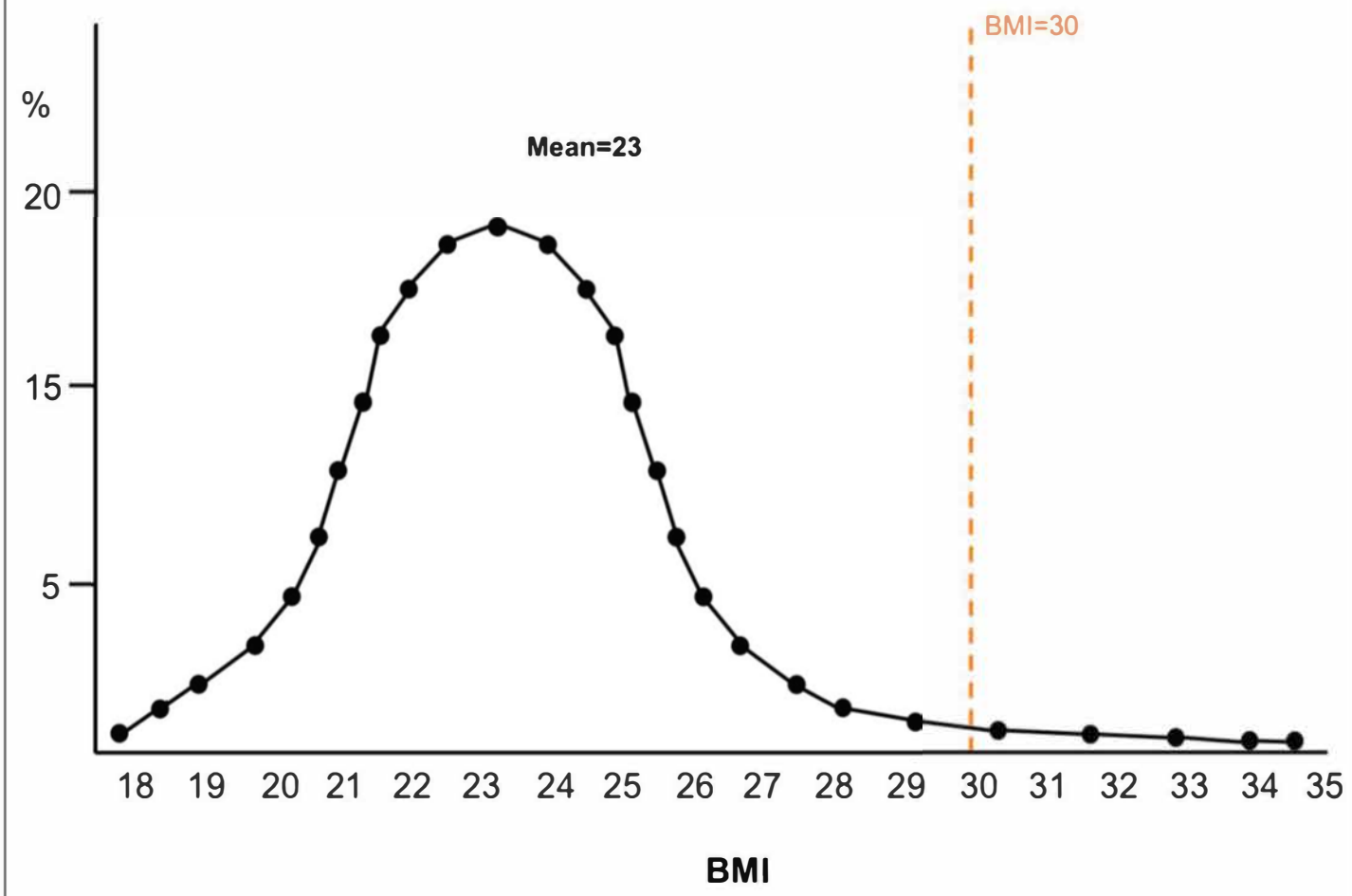


3. Four principles that can help

a. Population health data as continuous, not binary

Figure 1. Distribution of BMI in two populations illustrating health as a continuum in the population

Panel A



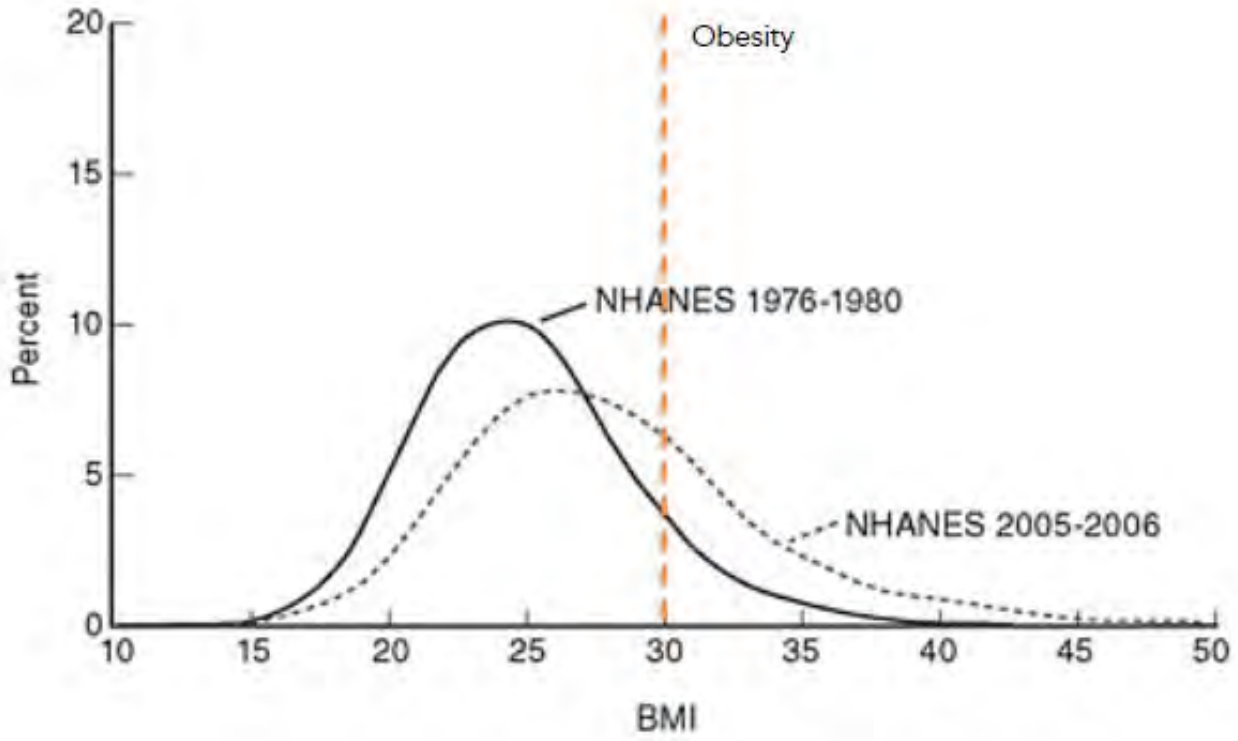


FIGURE 1-1 Changes in the distribution of body mass index (BMI) between 1976-1980 and 2005-2006 among U.S. adults aged 20-74. NOTE: NHANES = National Health and Nutrition Examination Survey, a continuous program of studies designed to assess the health and nutritional status of a nationally representative sample of children and adults in the United States. SOURCE: Ogden et al., 2007. https://www.nap.edu/read/12847/chapter/3#24

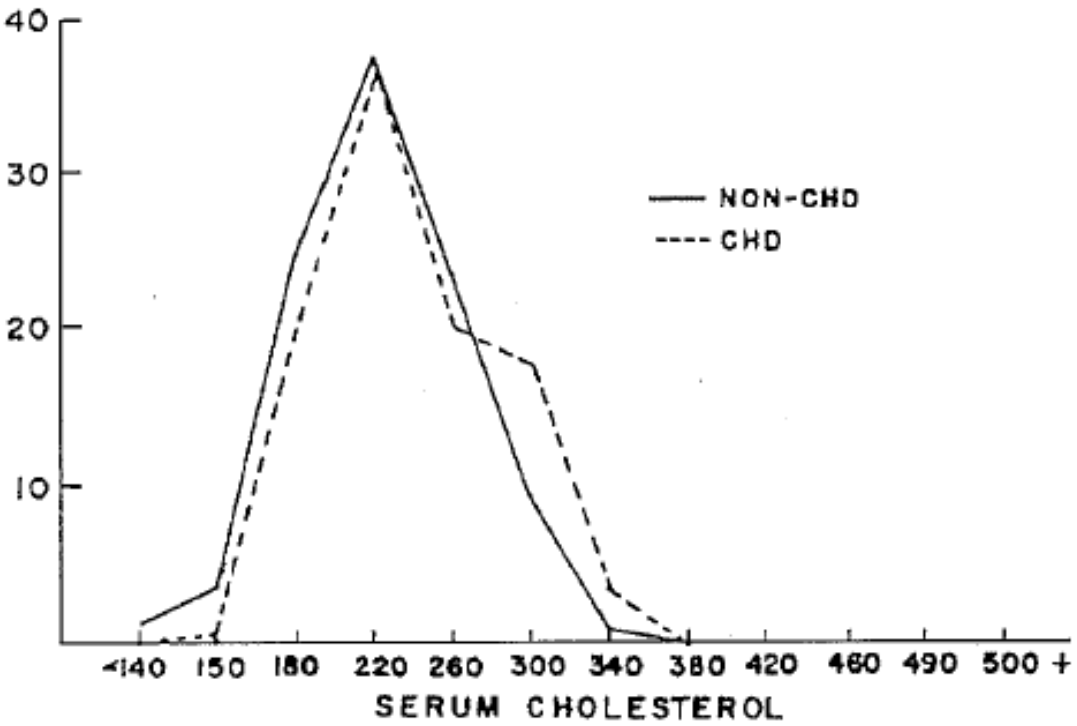
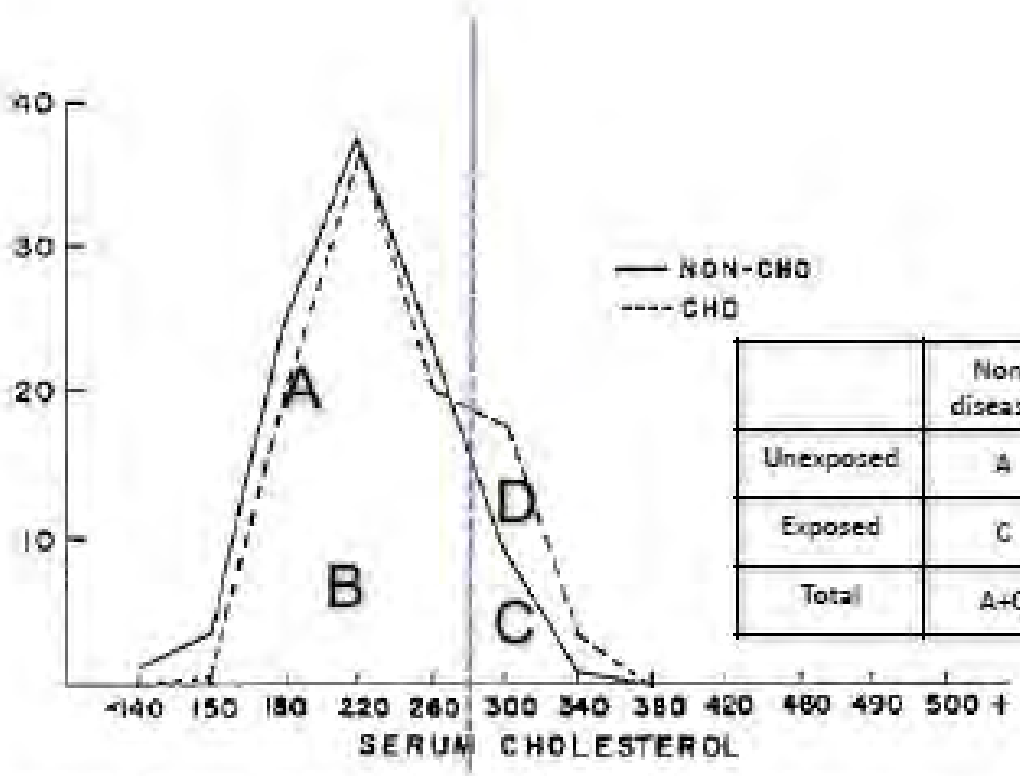


Figure 3 Percentage distribution of serum cholesterol levels (mg/dl) in men aged 50–62 who did or did not subsequently develop coronary heart disease (Framingham Study[5])

Rose G. Sick individuals and sick populations. International Journal of Epidemiology. 1985;14:32-38



	Non-diseased	Diseased	Total
Unexposed	A	B	A+B
Exposed	C	D	C+D
Total	A+C	B+D	N

Figure 3 Percentage distribution of serum cholesterol levels (mg/dl) in men aged 50–62 who did or did not subsequently develop coronary heart disease (Framingham Study[5])

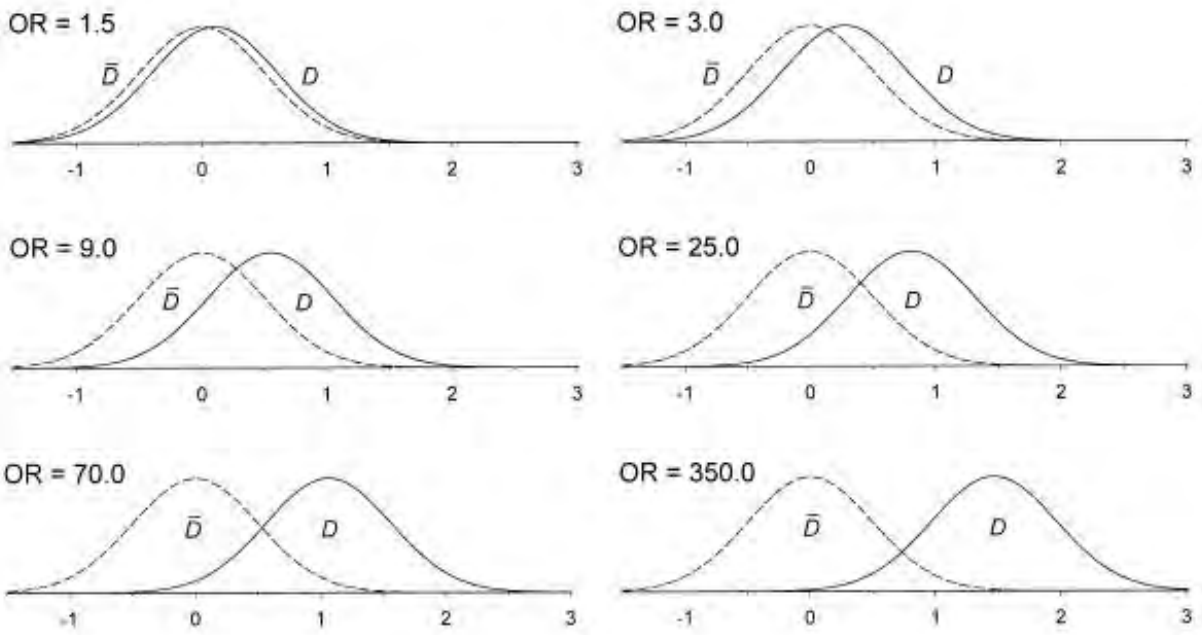
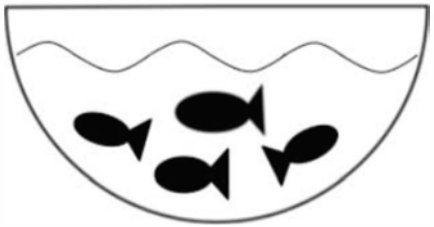


FIGURE 2. Probability distributions of a marker, X, in cases (solid curves) and controls (dashed curves) consistent with the logistic model $\text{logit}P(D = 1|X) = \alpha + \beta X$. It has been assumed that X has a mean of 0 and a standard deviation of 0.5 in controls so that a unit increase represents the difference between the 84th and 16th percentiles of X in controls. The marker is normally distributed, with the same variance in cases. The odds ratio (OR) per unit increase in X is shown.

Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology* 2004; 159:882-890.

b. Illuminating ubiquitous causes

Figure 1. A metaphor for ubiquity



The goldfish are surrounded by water and everything they do is influenced by the quality of the water in which they live; therefore, water is a ubiquitous factor influencing the fish and needs to be taken into consideration every time we may want to improve the lives of the fish.

Crack Babies: The Worst Threat Is Mom Herself

By Douglas J. Besharov

LAST WEEK in this city, Greater Southeast Community Hospital released a 7-week-old baby to her homeless, drug-addicted mother even though the child was at severe risk of pulmonary arrest. The hospital's explanation: "Because [the mother] demanded that the baby be released."

The hospital provided the mother with an apnea monitor to warn her if the baby stopped breathing while asleep, and trained her in CPR. But on the very first night, the mother went out drinking and left the child at a friend's house—without the monitor. Within seven hours, the baby was dead. Like Dooney Waters, the 6-year-old living in his mother's drug den, whose shocking story was reported in The Washington Post last week, this child was all but abandoned by ities.

September 17, 1989

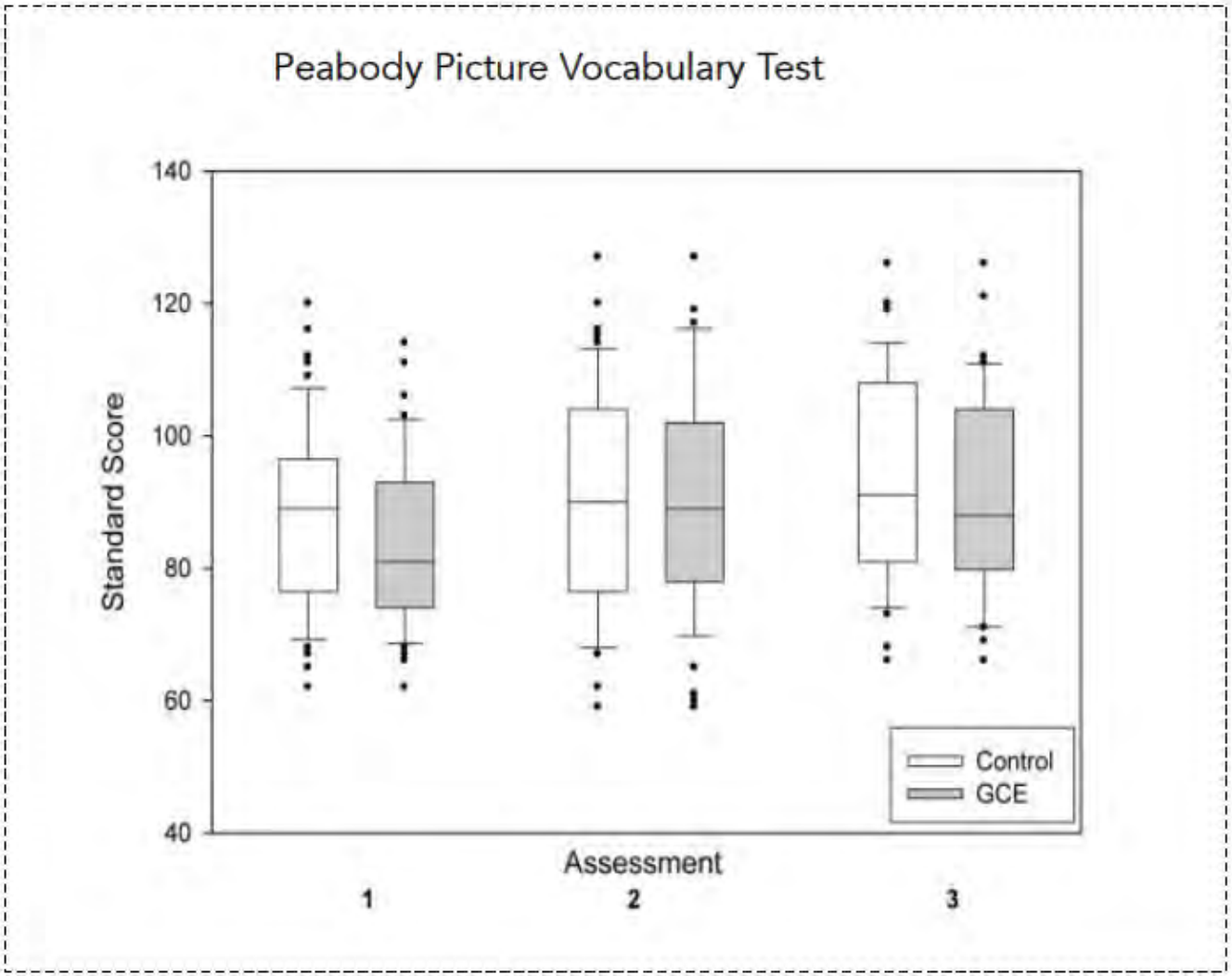
Crack's Toll Among Babies: A Joyless View

September 6, 1988

Cocaine: Litany of Fetal Risks Grows

CHILDREN OF COCAINE

(By Charles Krauthammer)

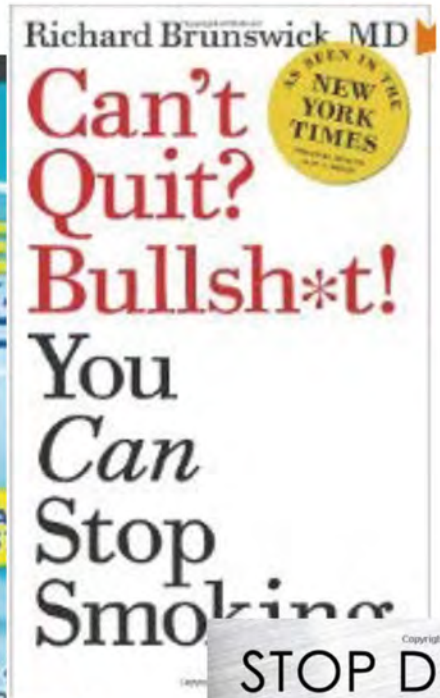


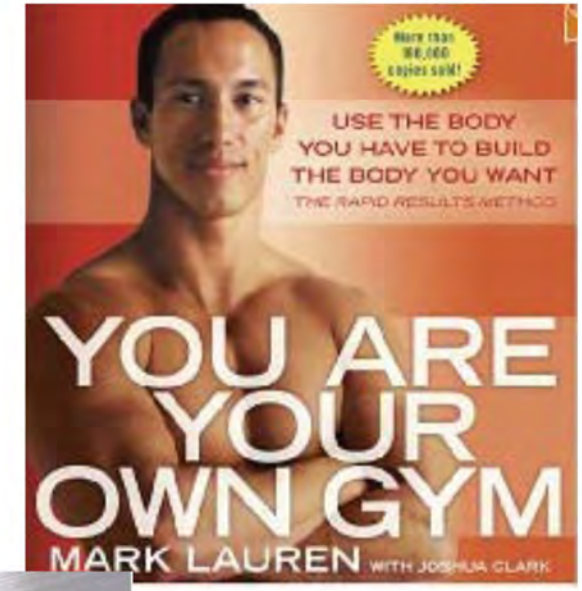
Betancourt LM et al. Adolescents with and without gestational cocaine exposure: Longitudinal analysis of inhibitory control, memory and receptive language. *Neurotoxicol Teratol* 2011; 33(1): 36-46.

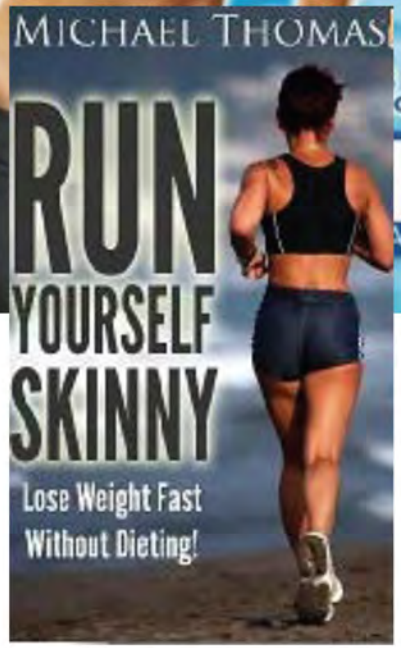
Predictor for Peabody Picture Vocabulary Test score	Coefficient	P-value
Gestational cocaine exposure	-2.89	0.26
Assessment no.	2.72	<0.001
Gestational cocaine exposure x assessment no.	0.58	0.51
Age at 1st assessment	-0.36	0.76
Female gender	-4.93	0.058
Parental nurturance	-0.31	0.89
Environmental stimulation	5.91	0.039
Caregiver BDI-II depression score	0.03	0.84

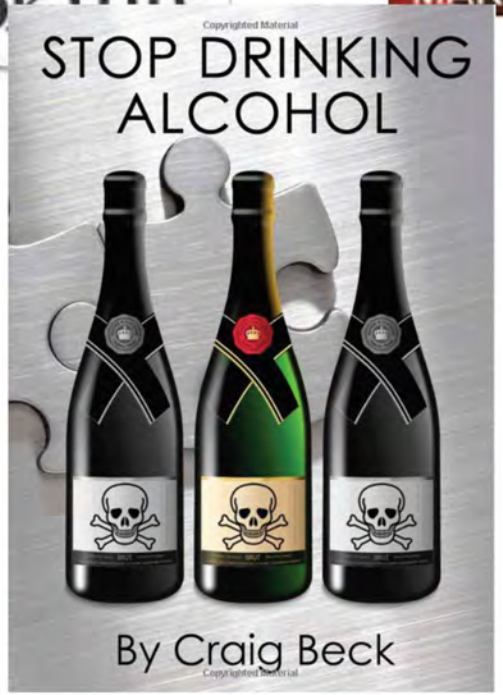
Betancourt LM et al. Adolescents with and without gestational cocaine exposure: Longitudinal analysis of inhibitory control, memory and receptive language. *Neurotoxicol Teratol* 2011; 33(1): 36-46.

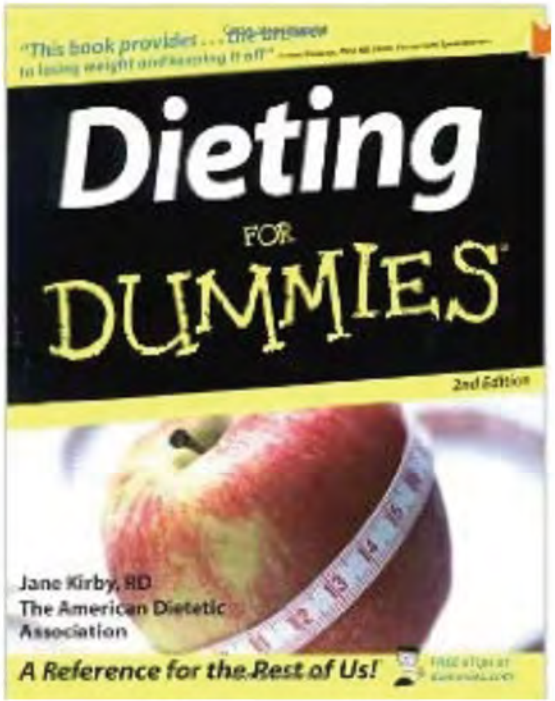

RIPPED ABS IN 6 WEEKS
6
WEE
SIX
PAC
FLAT A
FAS


Richard Brunswick, MD
AS SEEN IN THE NEW YORK TIMES
Can't Quit? Bullsh*t!
You Can Stop Smoking


USE THE BODY YOU HAVE TO BUILD THE BODY YOU WANT
YOU ARE YOUR OWN GYM
MARK LAUREN WITH JOSHUA CLARK


MICHAEL THOMAS
RUN YOURSELF SKINNY
Lose Weight Fast Without Dieting!

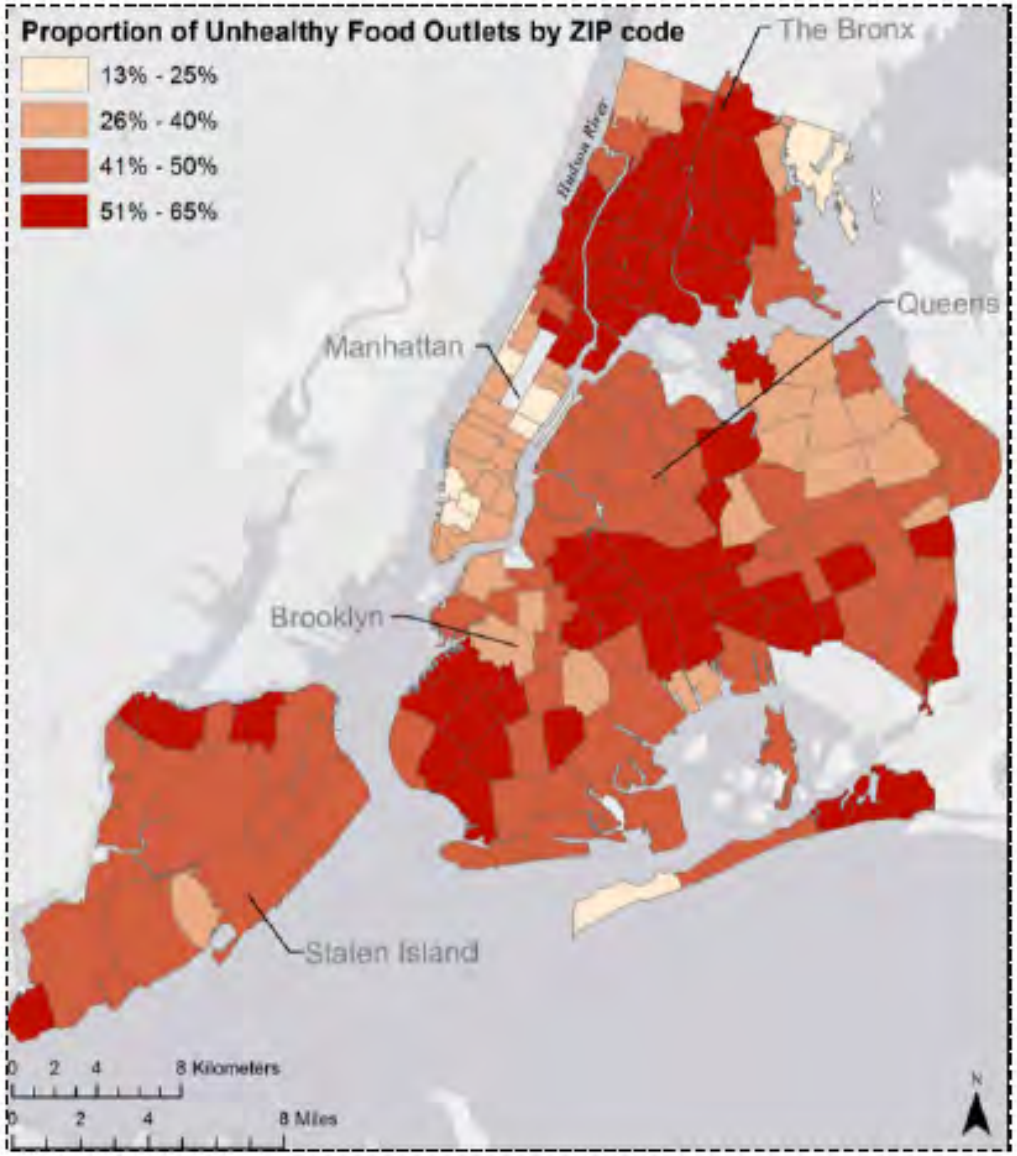

Copyrighted Material
STOP DRINKING ALCOHOL
By Craig Beck


Dieting
FOR
DUMMIES
2nd Edition
Jane Kirby, RD
The American Dietetic Association
A Reference for the Rest of Us!

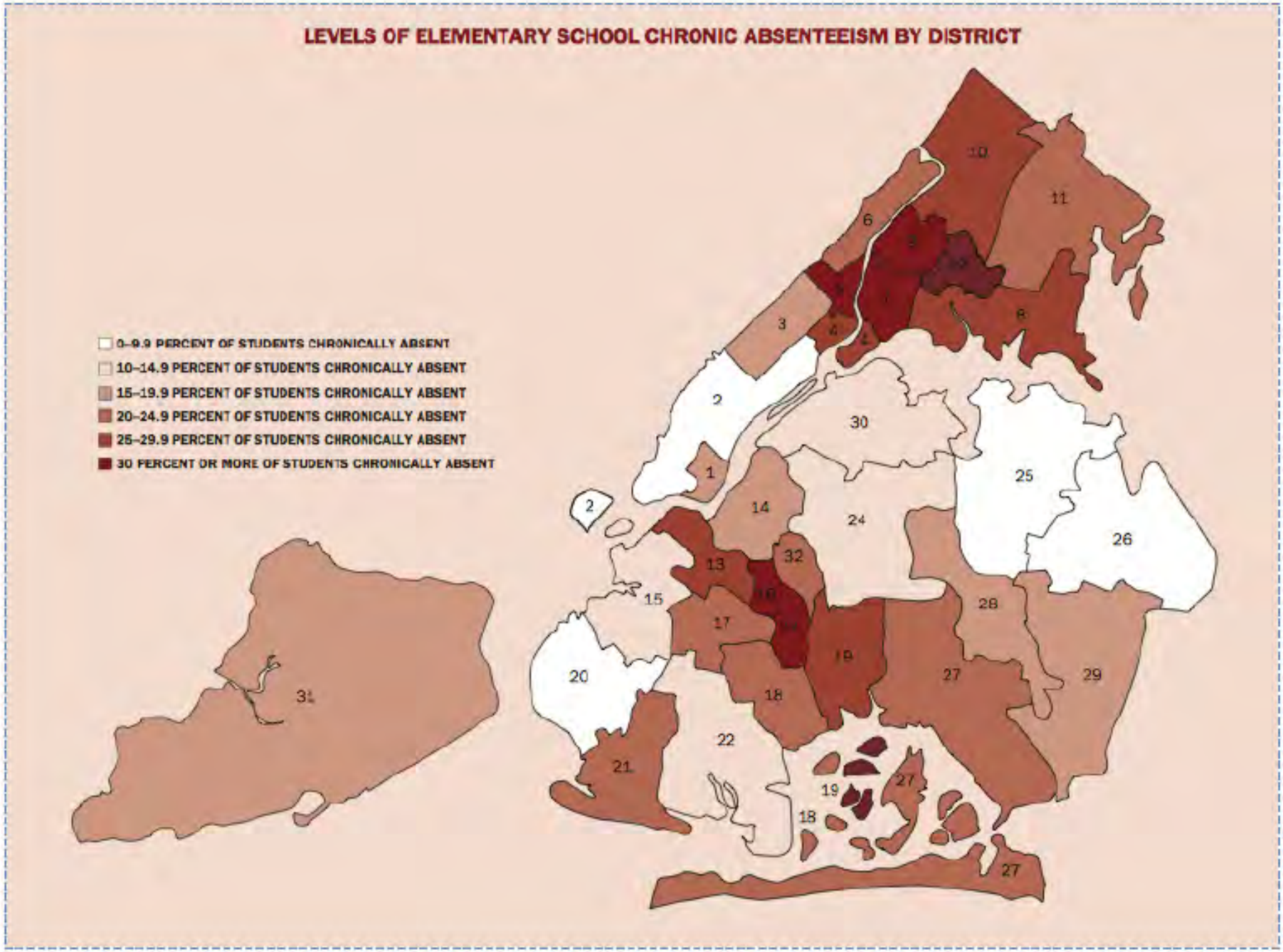
Food		20 Years Ago	Today
Bagel		140 calories (3" diameter)	350 calories (6" diameter)
Muffin		210 calories (1.5 oz)	500 calories (4 oz)
Cheeseburger		333 calories	590 calories
Pasta (Spaghetti & Meatballs)		500 calories	1025 calories
French Fries		210 calories (2.4 oz)	610 calories (6.9 oz)
Soda		85 calories (6.5 oz)	250 calories (20 oz)
Theater Popcorn		270 calories (5 cups)	630 calories (1 tub)
Turkey Sandwich		320 calories	820 calories
Pizza		500 calories (2 slices)	850 calories (2 calories)

http://www.yourweightmatters.org/portion-sizes-changed-time/

Poor food environment in New York City



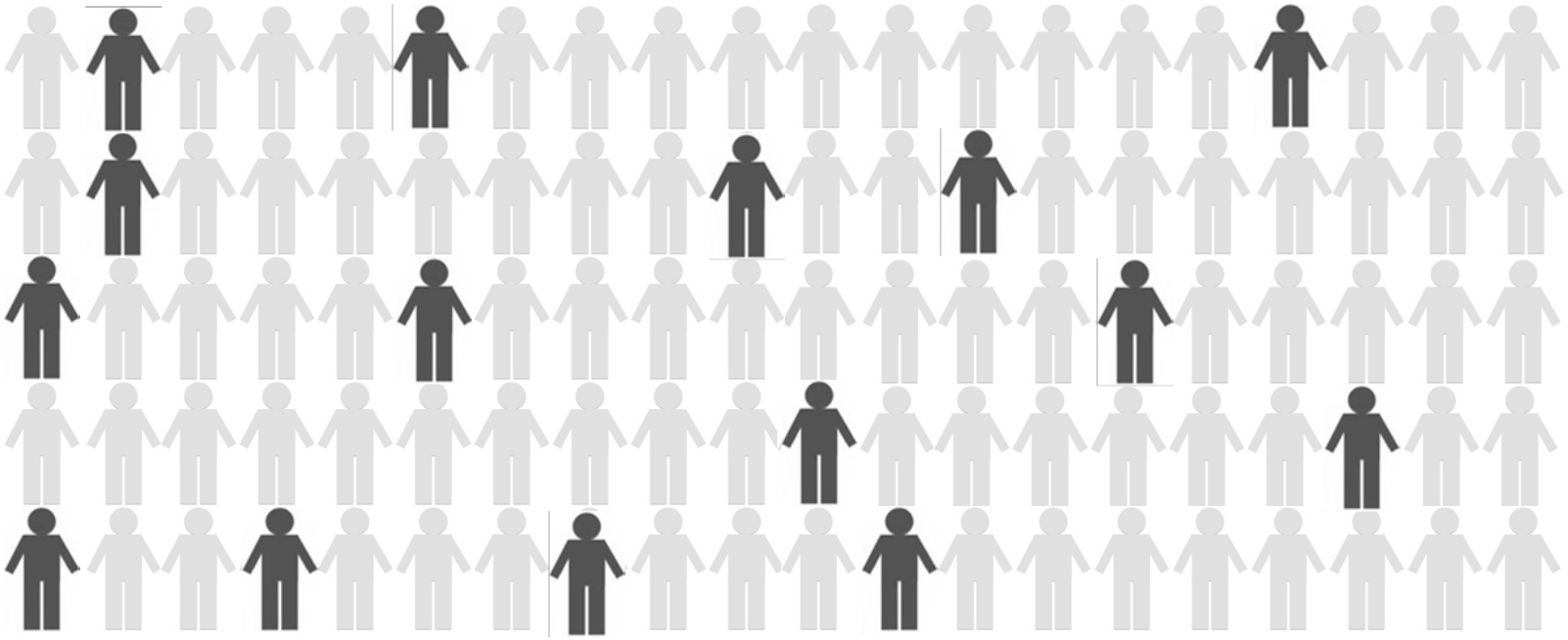
Stark J et al. Neighbourhood food environments and body mass index among New York City adults. *J Epidemiol Community Health* 2013; 67: 736-742



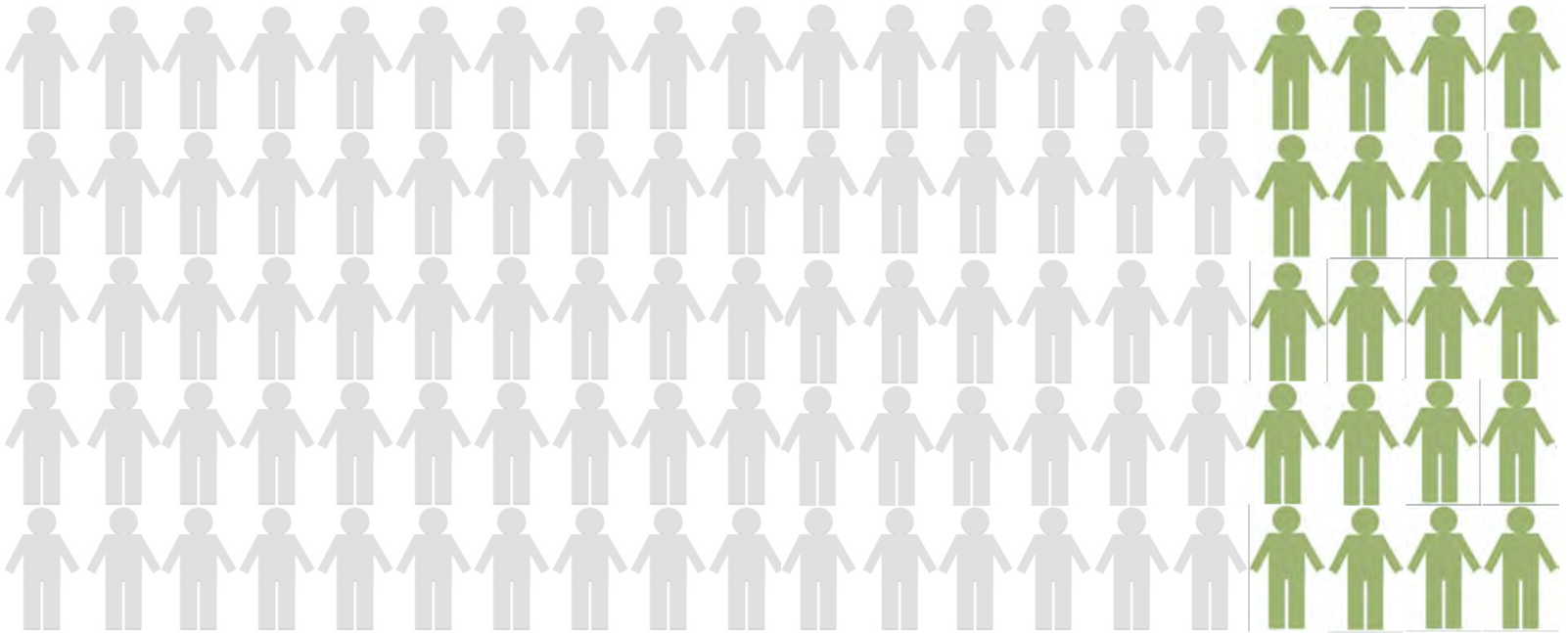
http://www.usnews.com/cmsmedia/13/cf/75a042c94c0fb078efa0790b2e73/141110-hechingerabsenteeism-graphic.absenteeism.png

c. The role of co-occurring causes

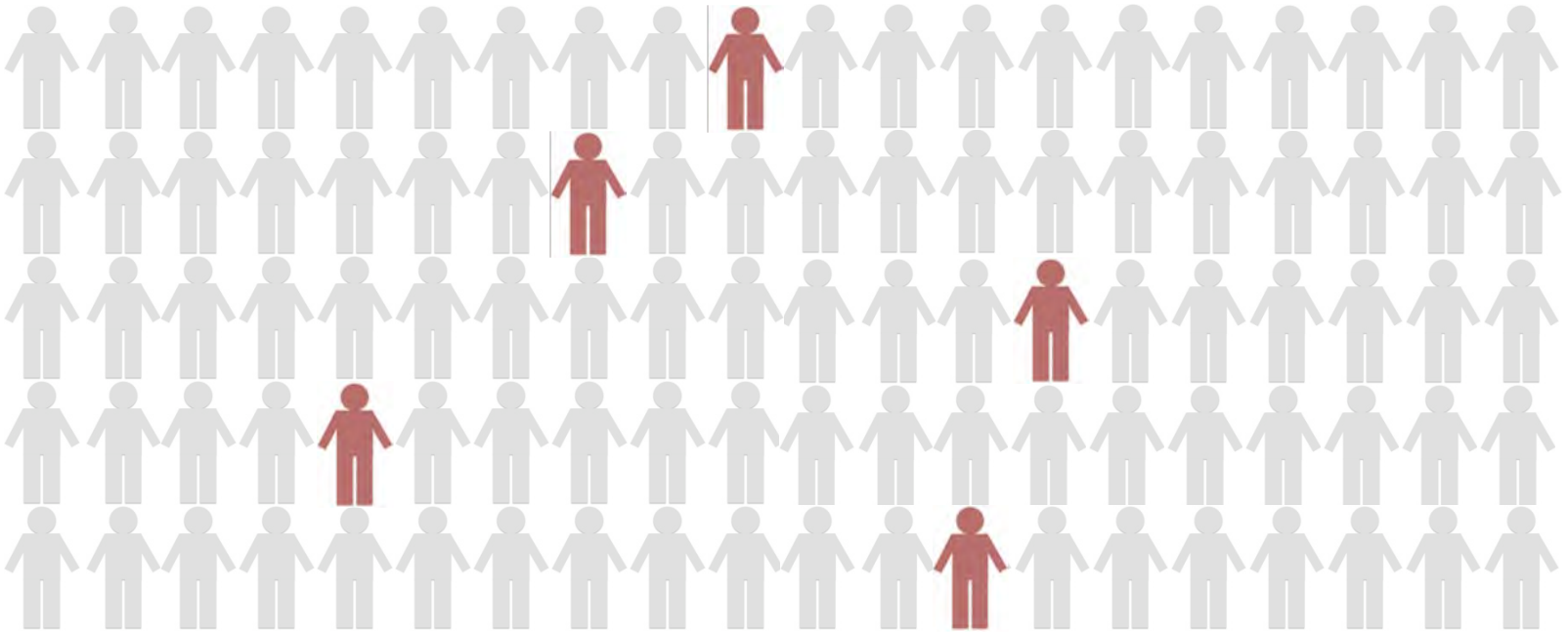
How much of our obesity risk is determined by our genes?



= GE+

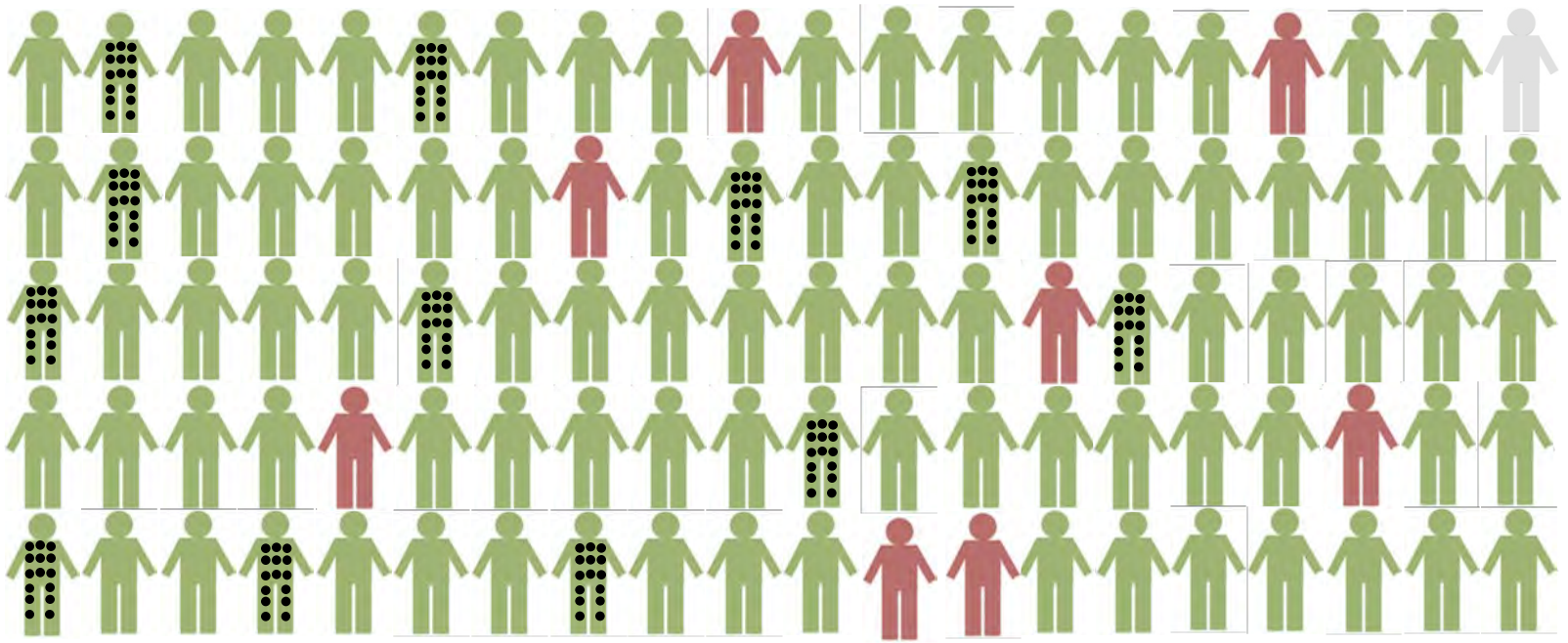


= ENV+



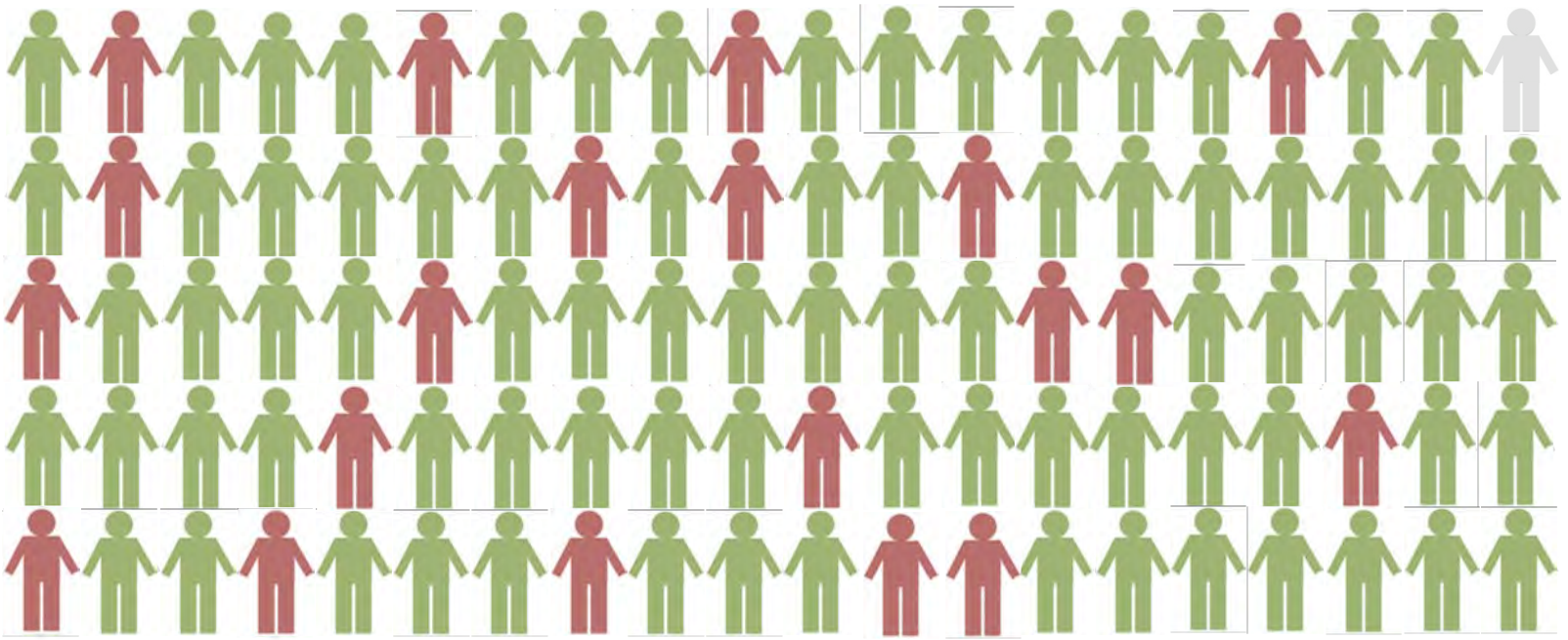
= OB+

Scenario 1



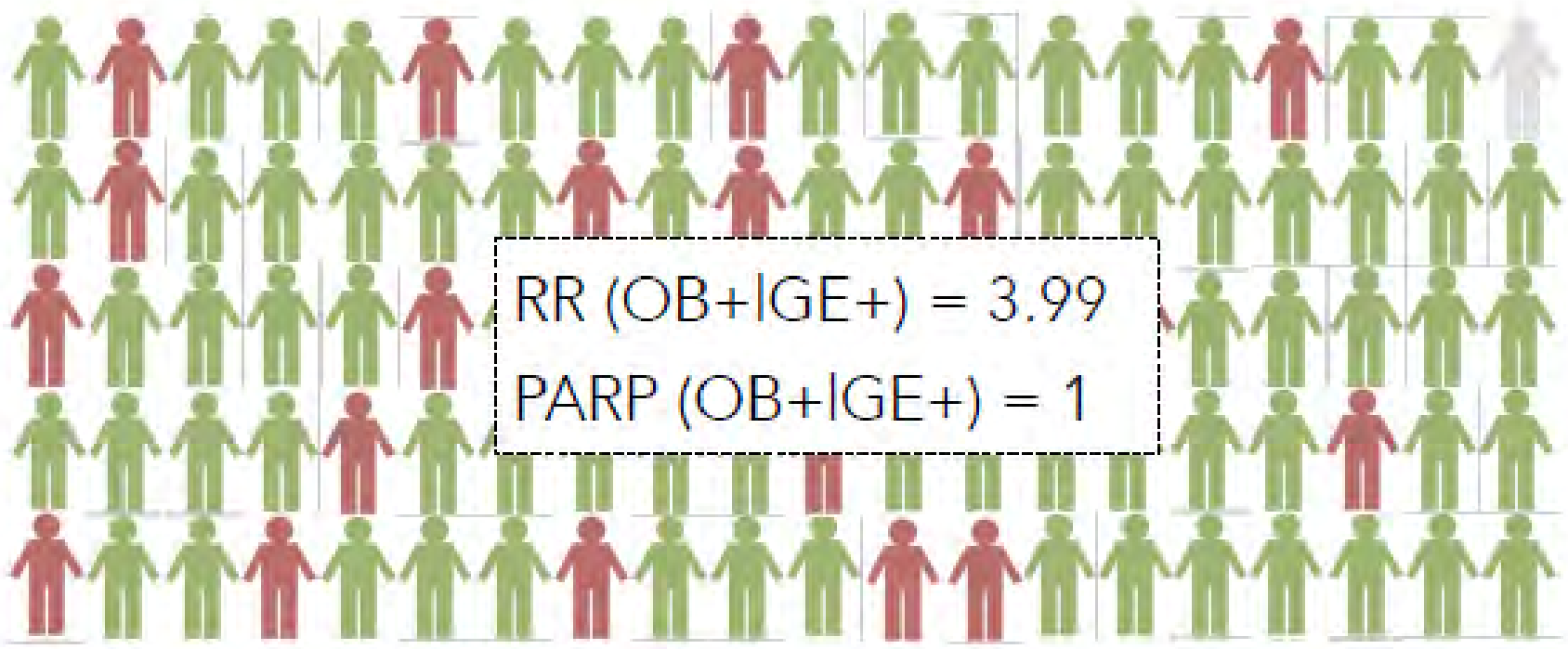
= GE+ = OB+ = ENV+

Scenario 1



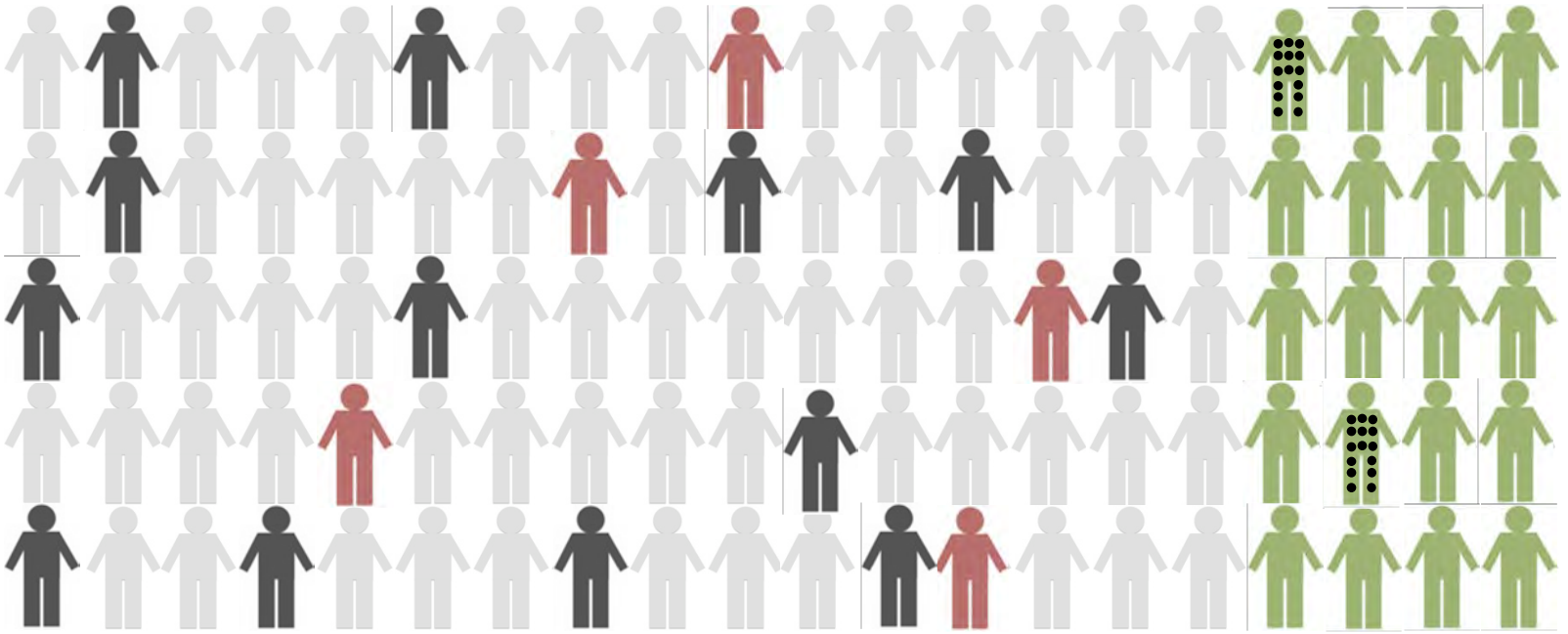
= GE+ = OB+ = ENV+

Scenario 1



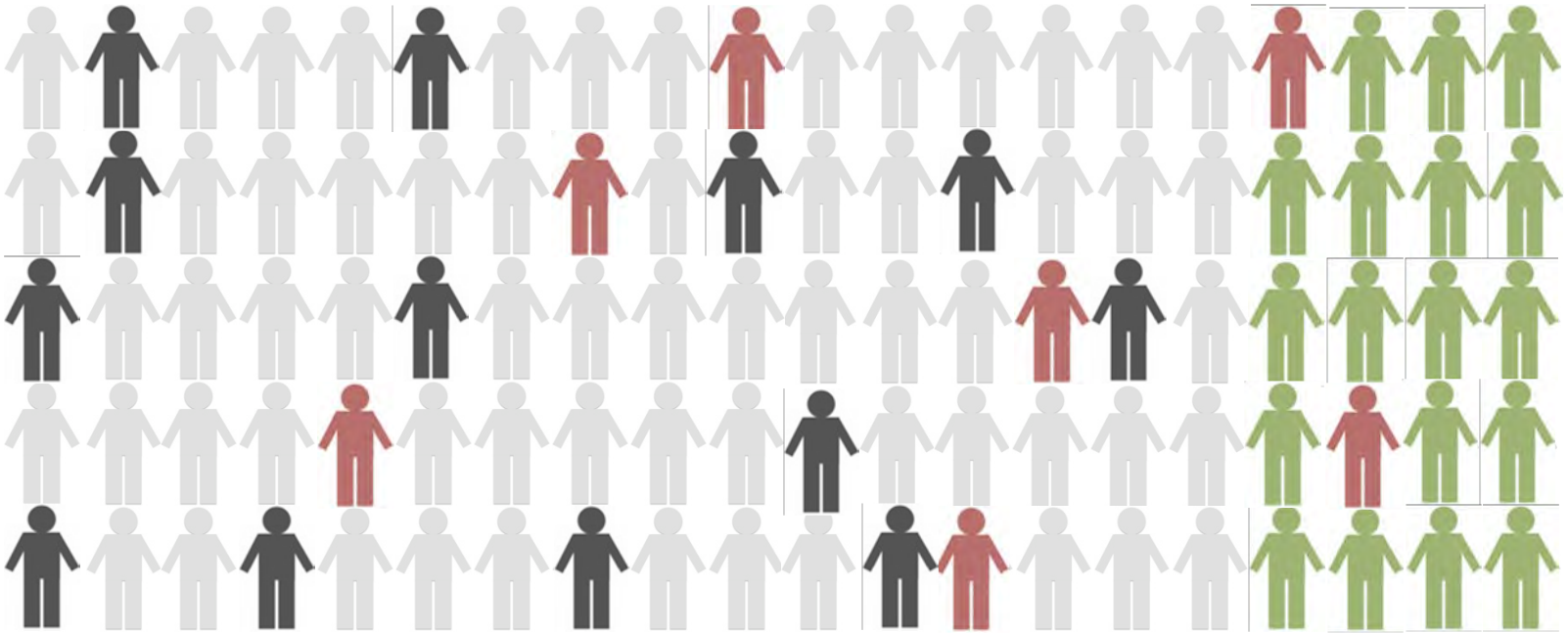
= GE+ = CA+ = ENV+

Scenario 2



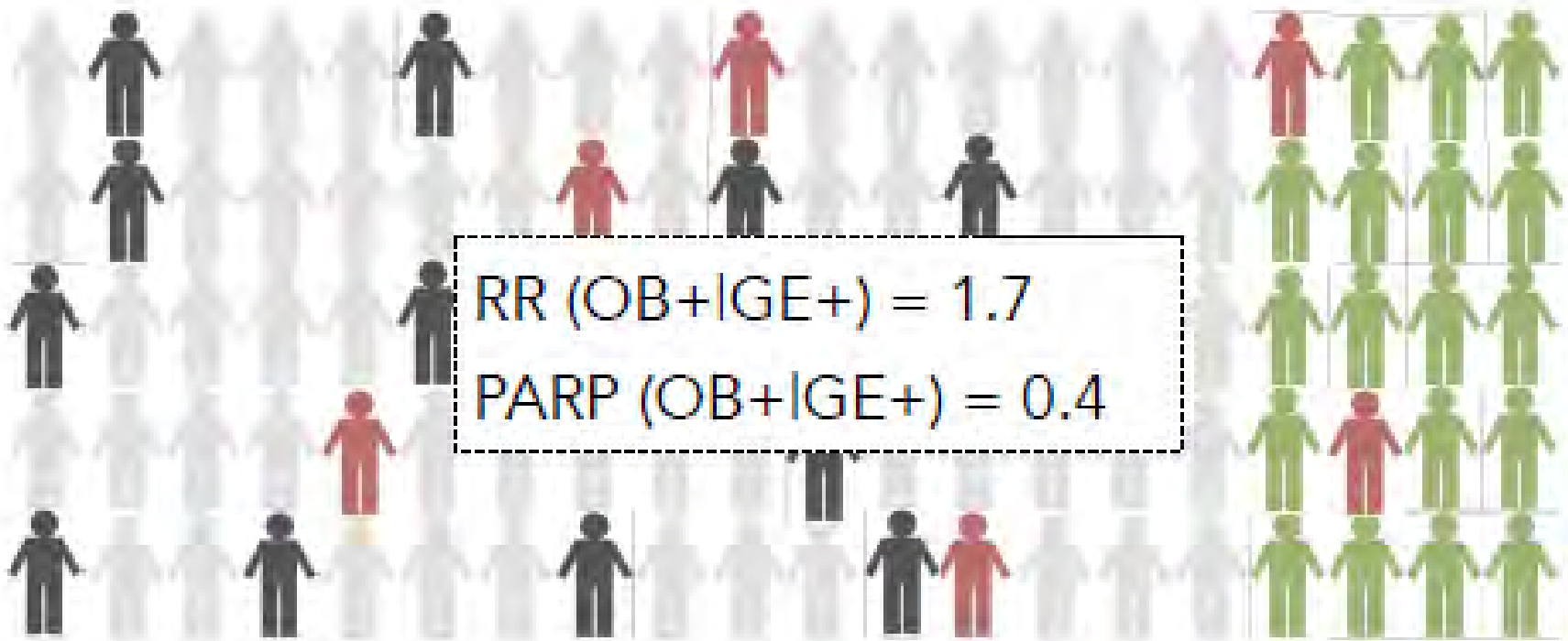
= GE+ = OB+ = ENV+

Scenario 2



= GE+ = OB+ = ENV+

Scenario 2



= GE+ = OB+ = ENV+

Therefore under a very plausible assumption of co-occurring causes, the gene-obesity association can only be understood if we understand the urban factors that create the conditions for disease

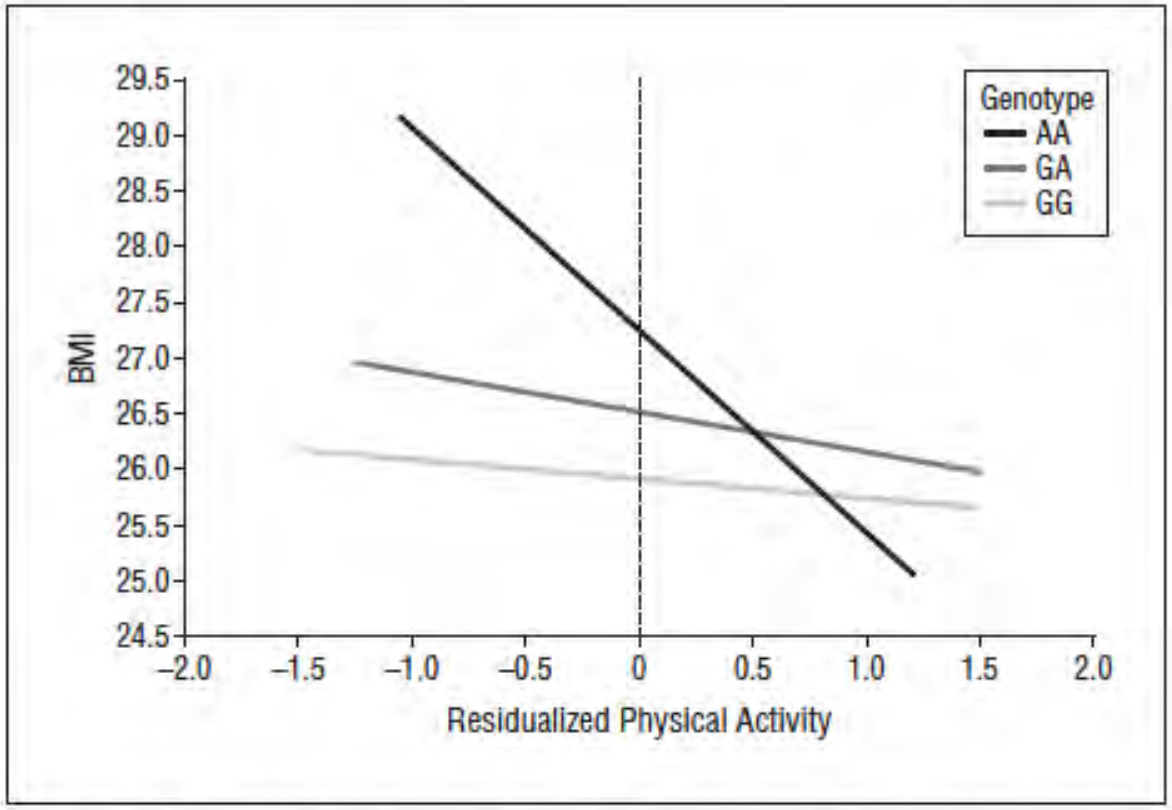
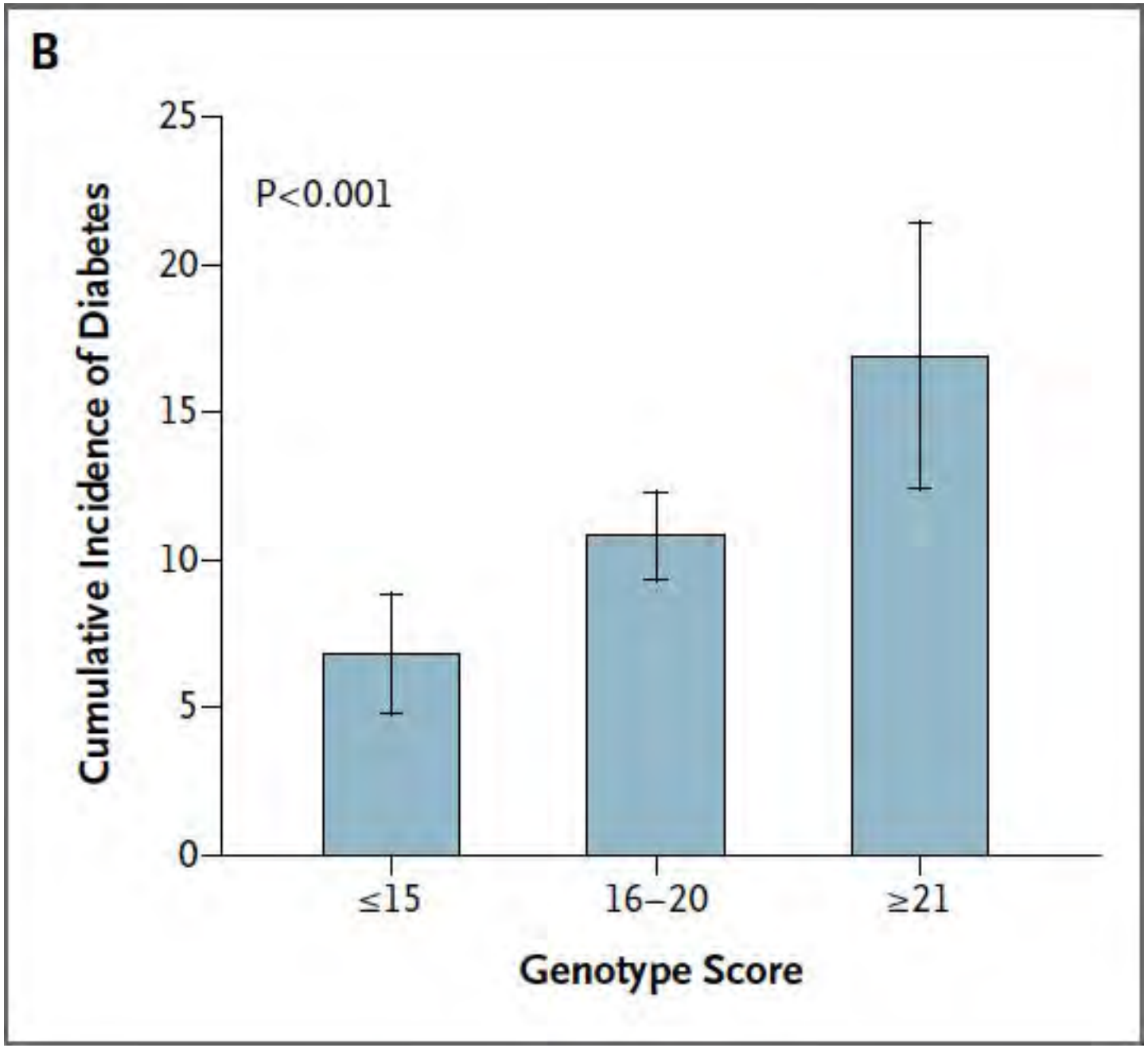
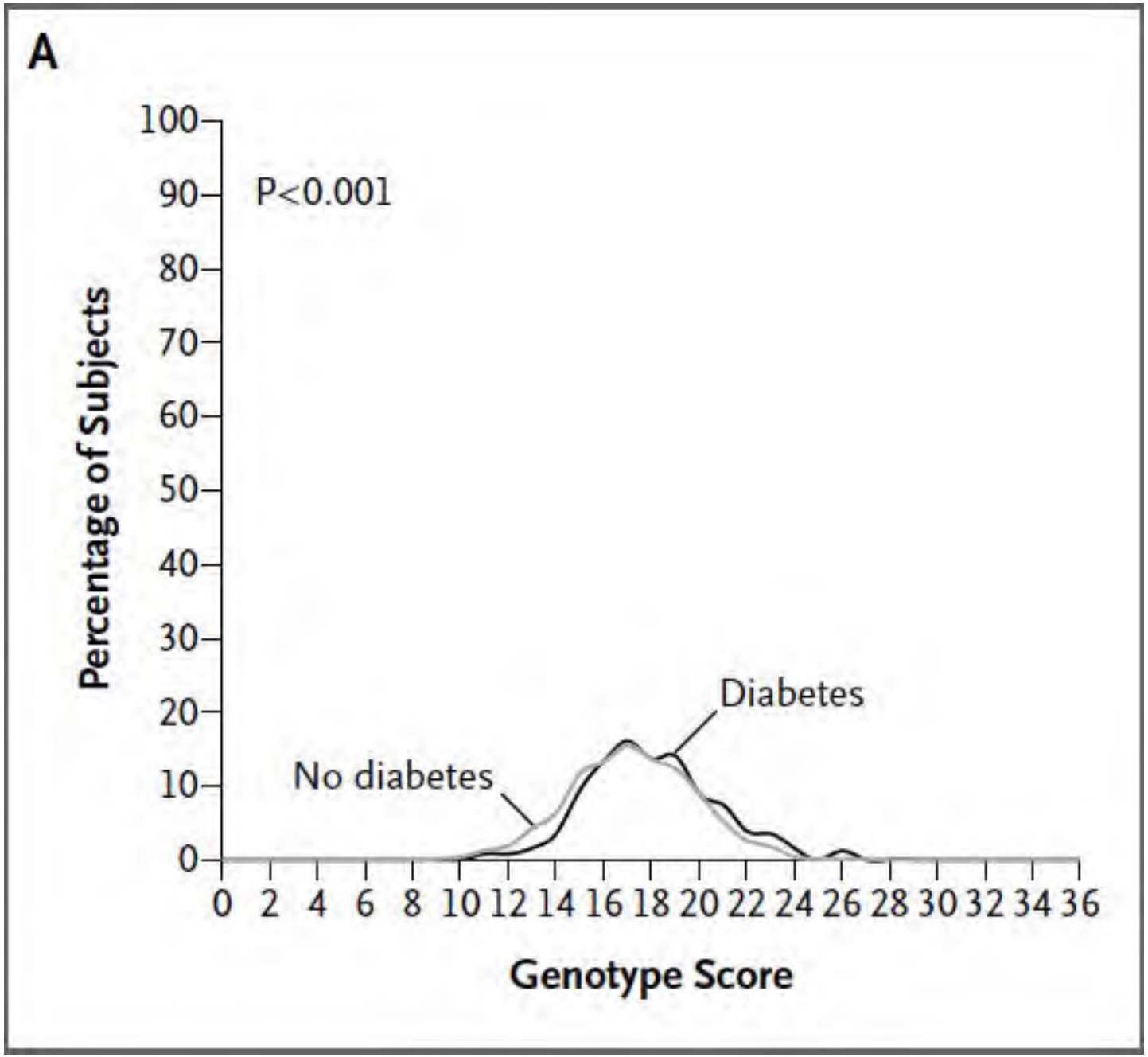


Figure 3. Predicted body mass index (BMI), calculated as weight in kilograms divided by height in meters squared, as a function of residualized age- and sex-specific ln-transformed physical activity accelerometer counts according to *FTO* rs1861868 genotypes. On the left side of the plot (low physical activity), BMI levels are strikingly dissimilar between rs1861868 genotypes. In contrast, on the right side of the plot, similar BMI levels can be seen across genotypes, particularly in subjects with very high levels of physical activity.

Rampersaud E, Mitchell BD, Pollin TI, Fu M, Shen H, O'Connell JR, et al. Physical Activity and the Association of Common FTO Gene Variants With Body Mass Index and Obesity. *Arch Intern Med.* 2008;168(16):1791-1797.



Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, Manning AK, Florez JC, Wilson PW, D'Agostino RB Sr, Cupples LA. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med. 2008 Nov 20;359(21):2208-19



Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, Manning AK, Florez JC, Wilson PW, D'Agostino RB Sr, Cupples LA. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med. 2008 Nov 20;359(21):2208-19

d. The centrality of health equity

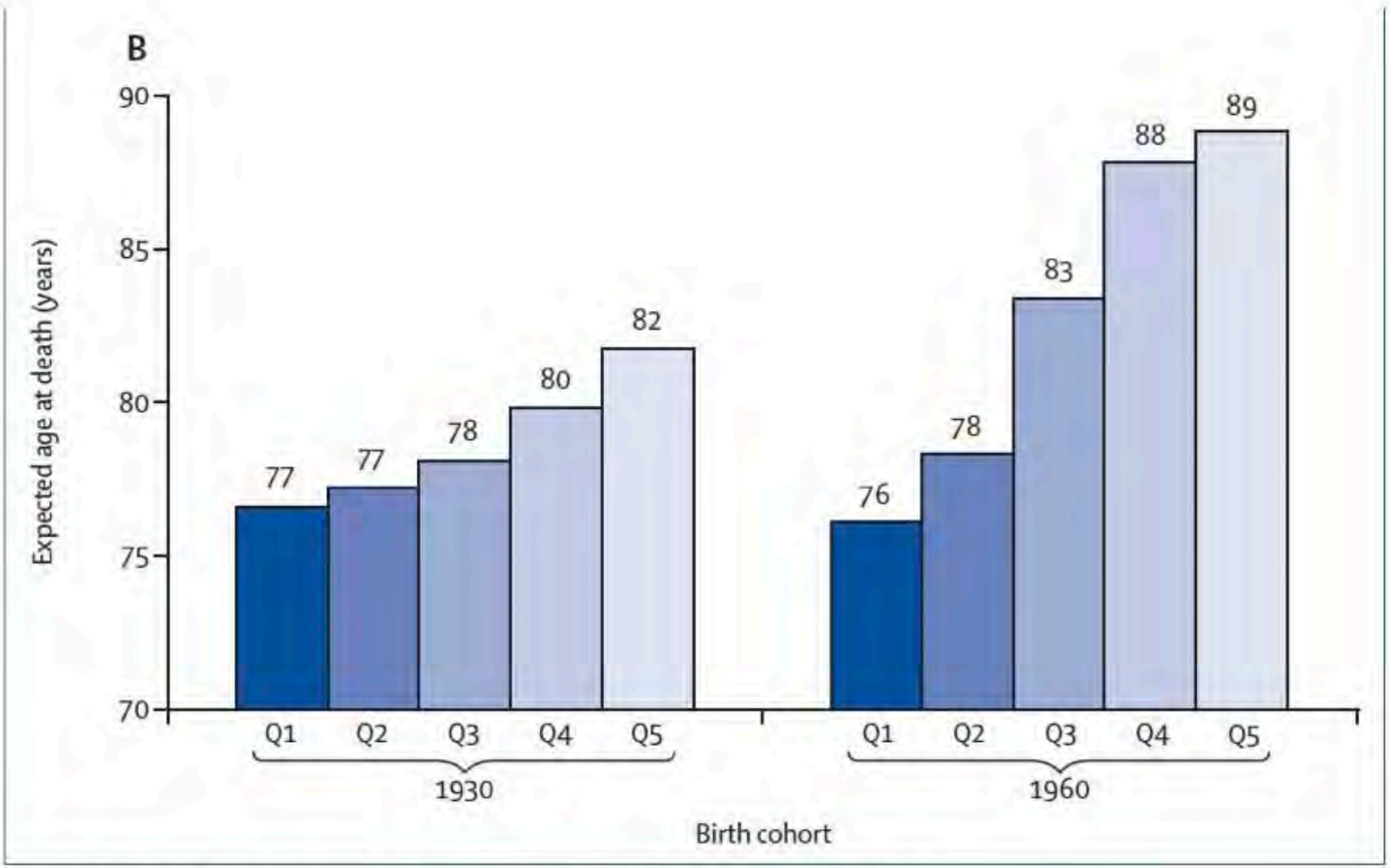
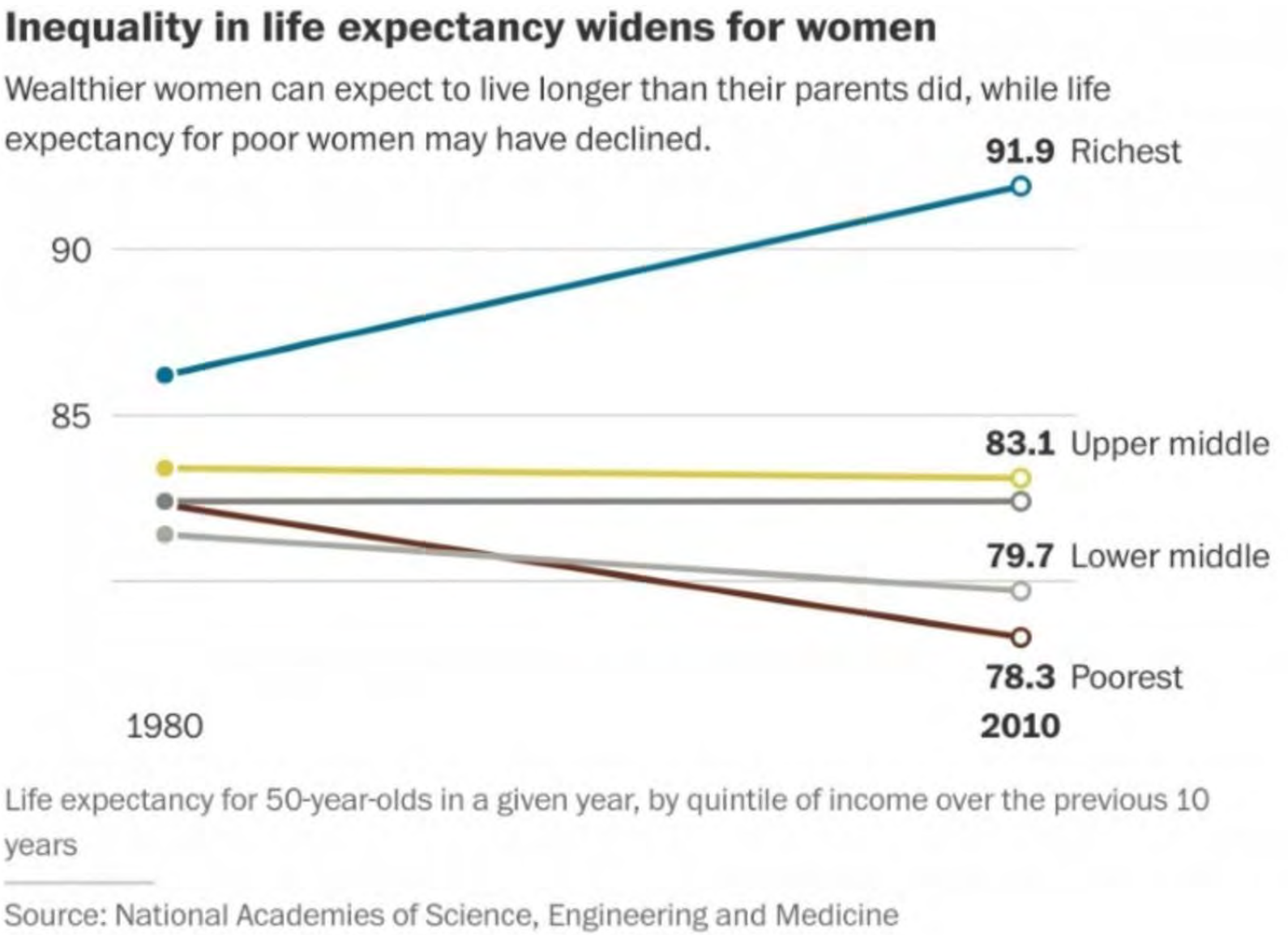


Figure 5: **Widening income-related inequalities in survival across birth cohorts**

Bor J, Cohen G, Galea S. Population Health in an Era of Rising Income Inequality: United States, 1980-2015. The Lancet. 2017; 389(10077):1475-1490



National Academy of Medicine. The growing gap in life expectancy by income: Implications for federal programs and policy responses. 2015. http://www.nap.edu/download.php?record_id=19015#

Figure 1. Gaining overall population health while increasing health inequity

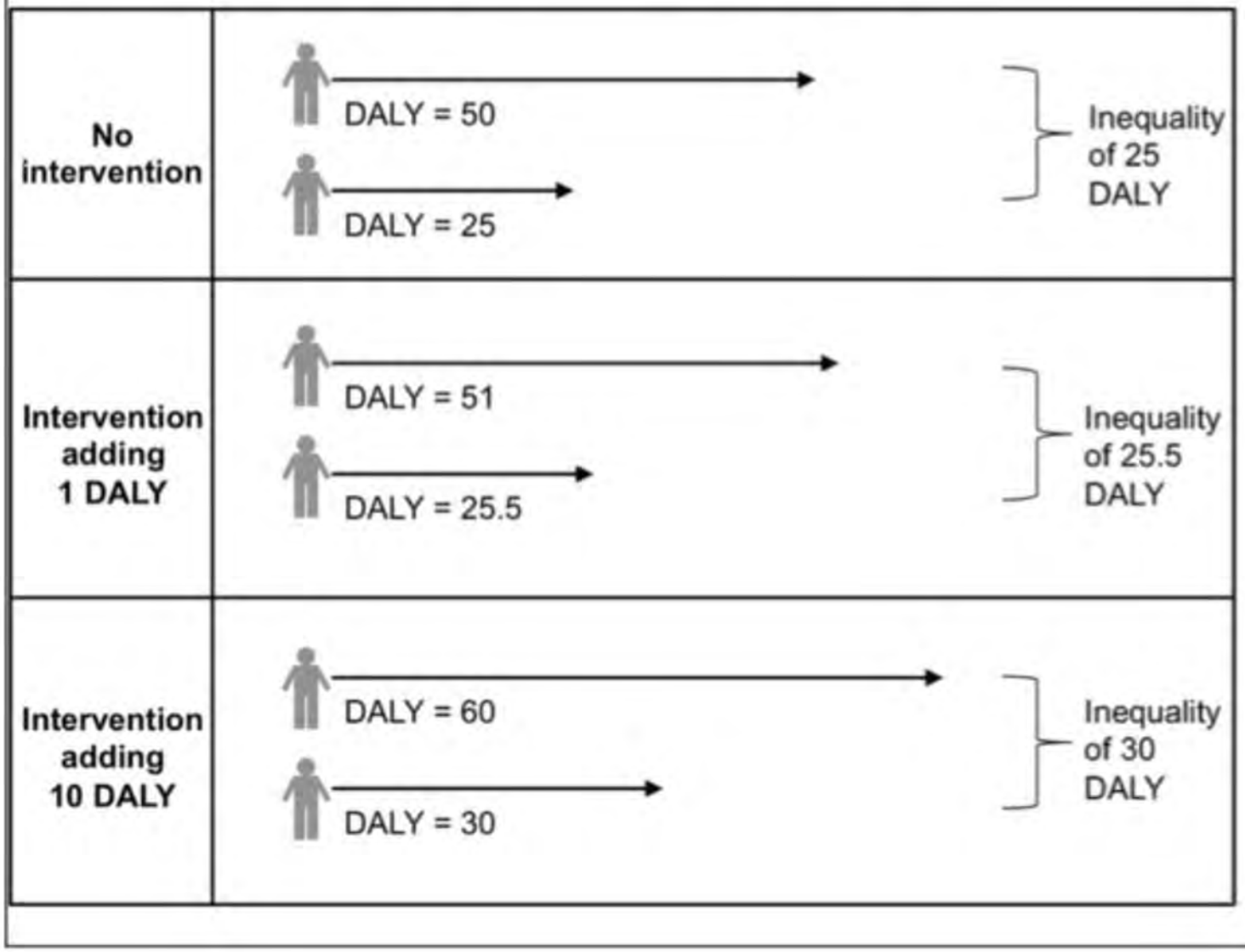
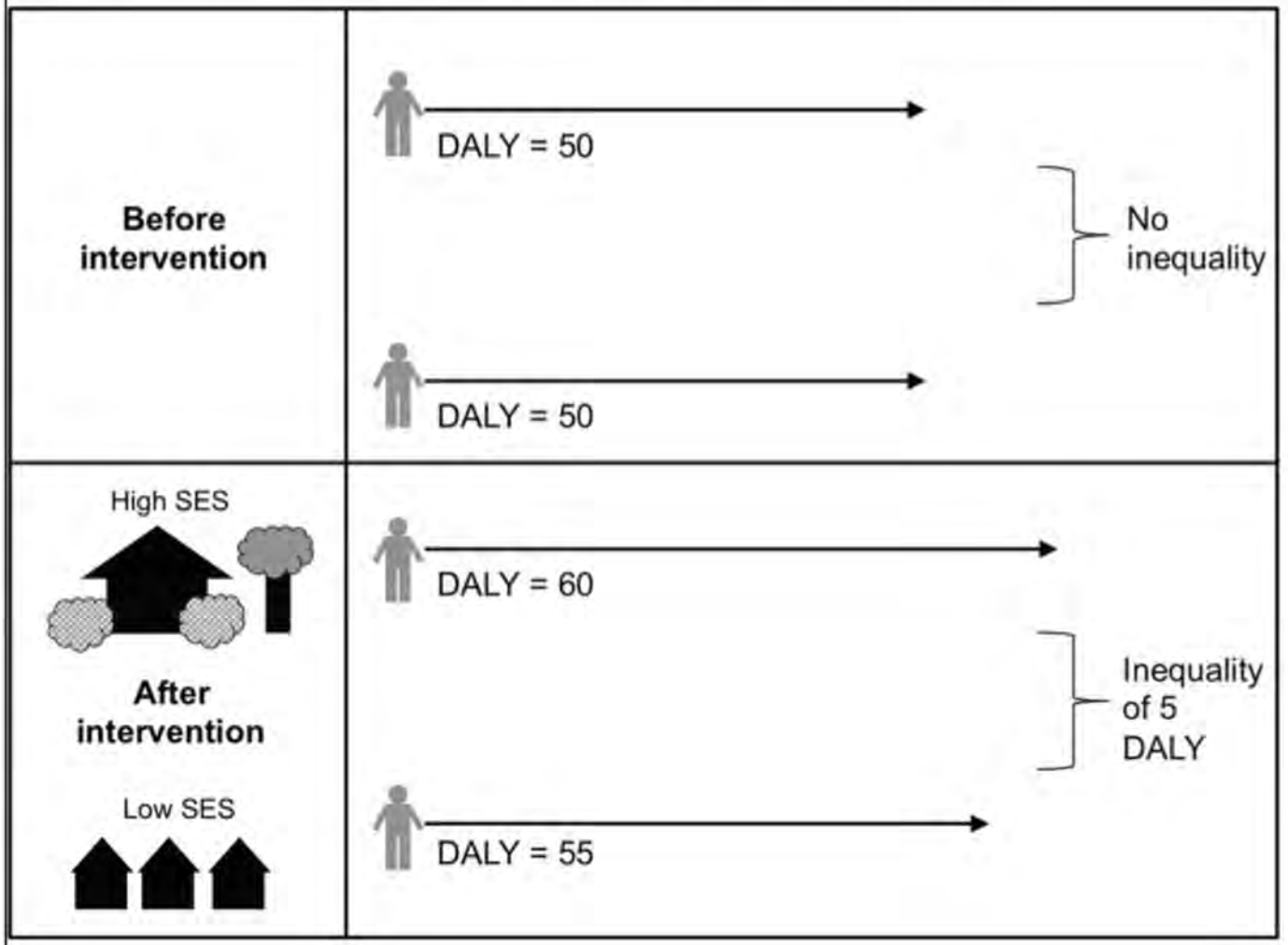
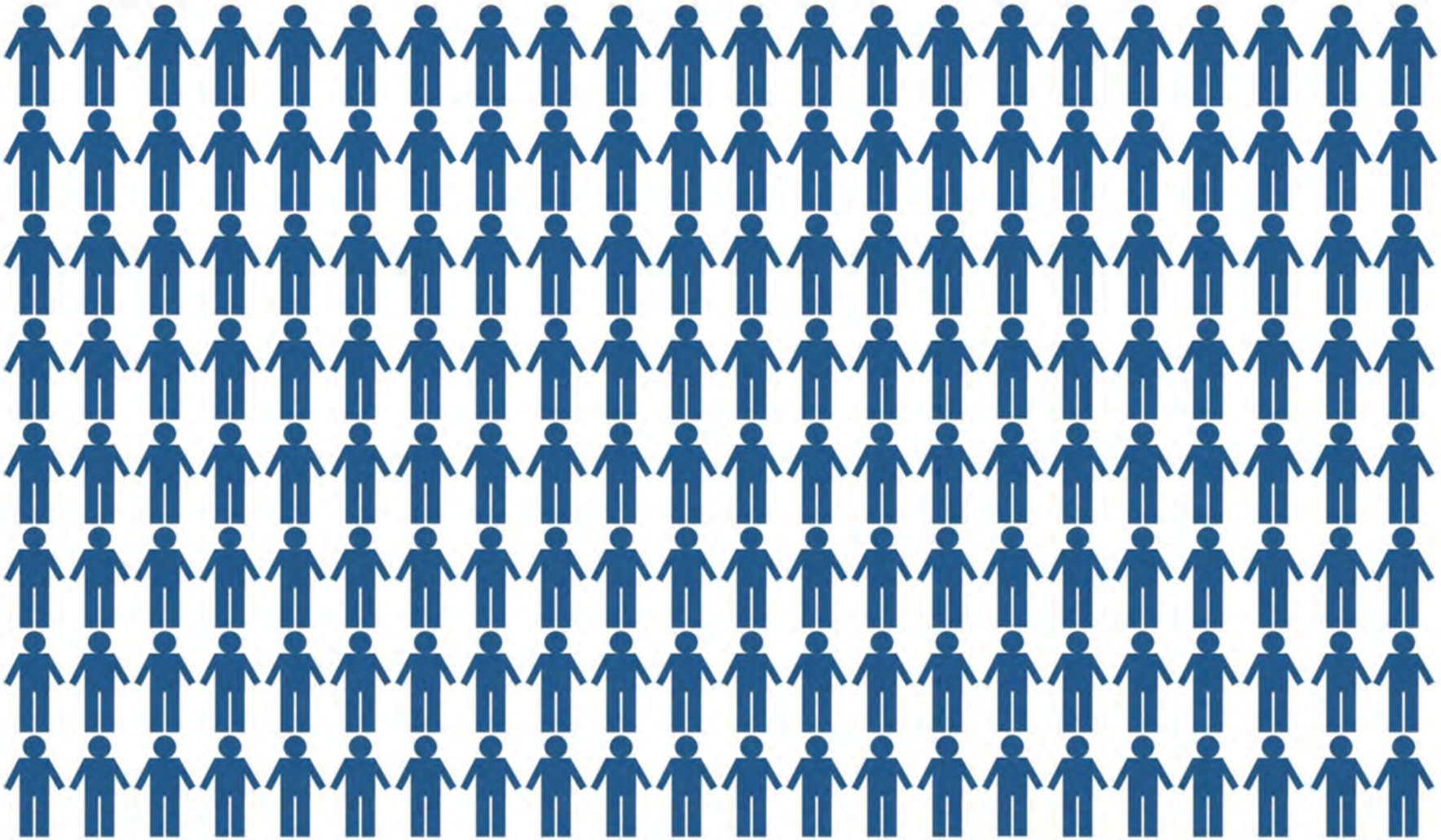


Figure 2. Gaining overall population health while creating health inequalities



4. Methods for population health



Populations are

1. Heterogeneous, ie have diversity of agents
2. Characterized by nonlinear dynamics
3. Characterized by contact structure, networks, organization
4. Have feedback, adaptation, learning, evolution
5. Stochastic with important tails
6. Display emergent properties

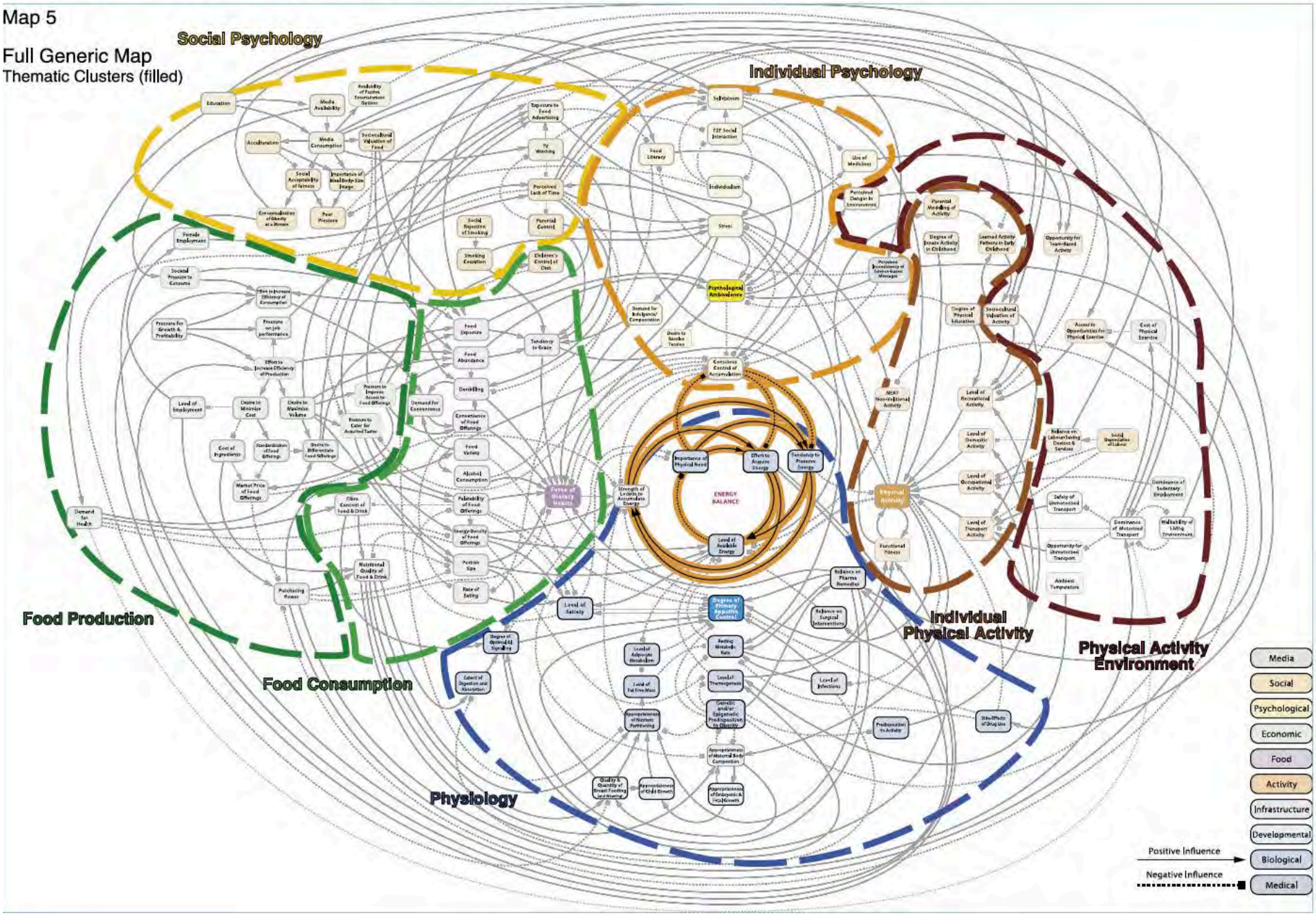
http://i.telegraph.co.uk/multimedia/archive/00796/crowded-britain_796405c.jpg

Populations are

1. Heterogeneous, ie have diversity of agents
2. Characterized by nonlinear dynamics
3. Characterized by contact structure, networks, organization
4. Have feedback, adaptation, learning, evolution
5. Stochastic with important tails
6. Display emergent properties

Complex systems

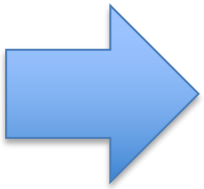
Figure 5.2: The full obesity system map with thematic clusters (see main text 5.1.2 for discussion)[17,18] Variables are represented by boxes, positive causal relationships are represented by solid arrows and negative relationships by dotted lines. The central engine is highlighted in orange at the centre of the map.



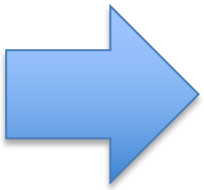
<http://www.foresight.gov.uk/OurWork/ActiveProjects/Obesity/KeyInfo/Index.asp>

Causes and counterfactuals

Observed

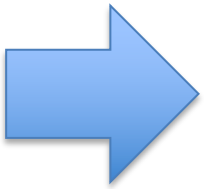


Counterfactual (parallel universe)

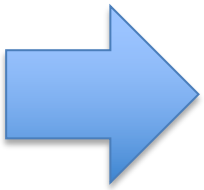


Causes and counterfactuals

Observed

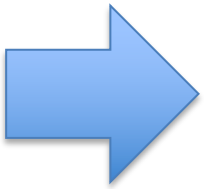


Counterfactual (parallel universe)

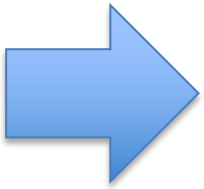


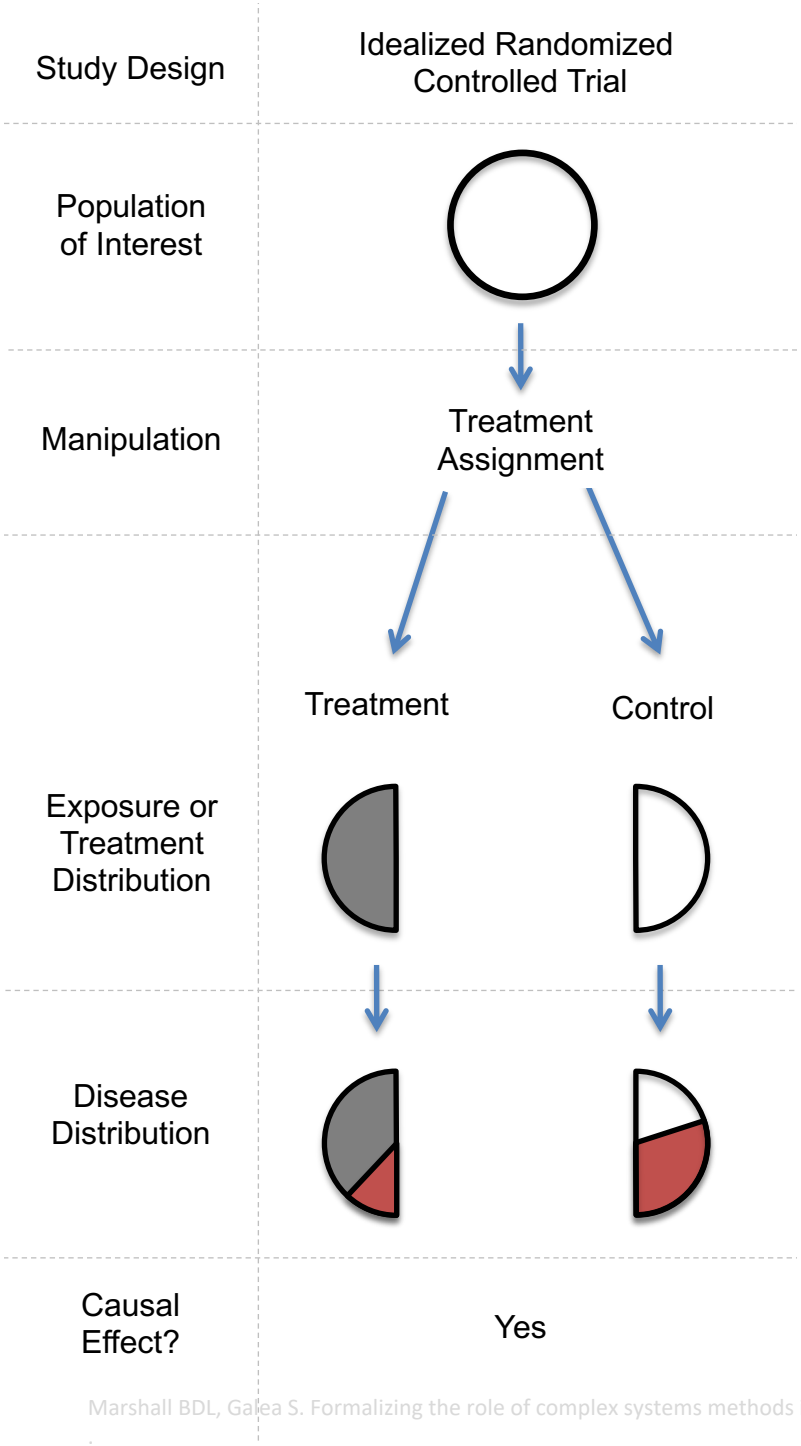
Causes and counterfactuals

Observed

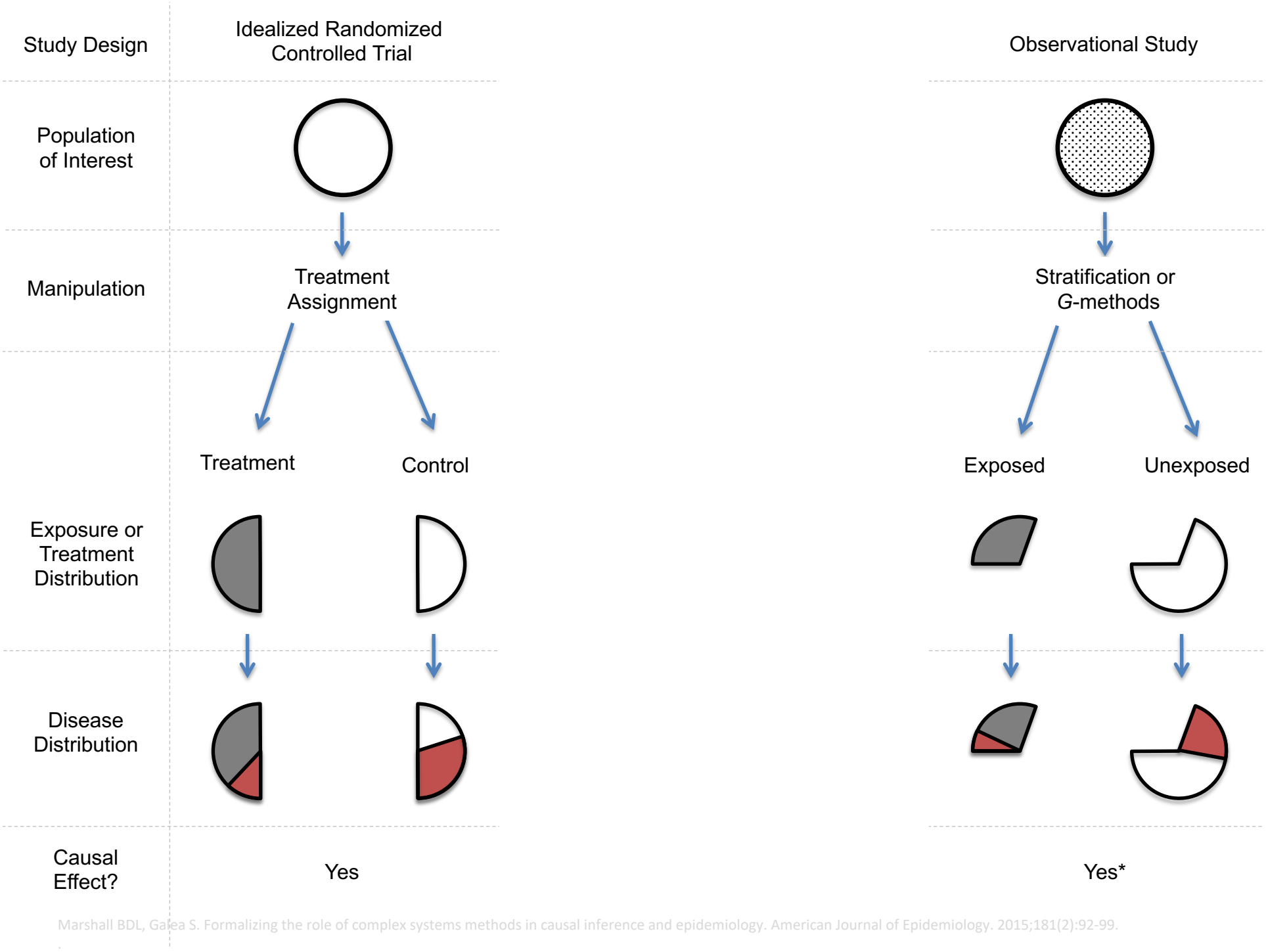


Counterfactual (parallel universe)

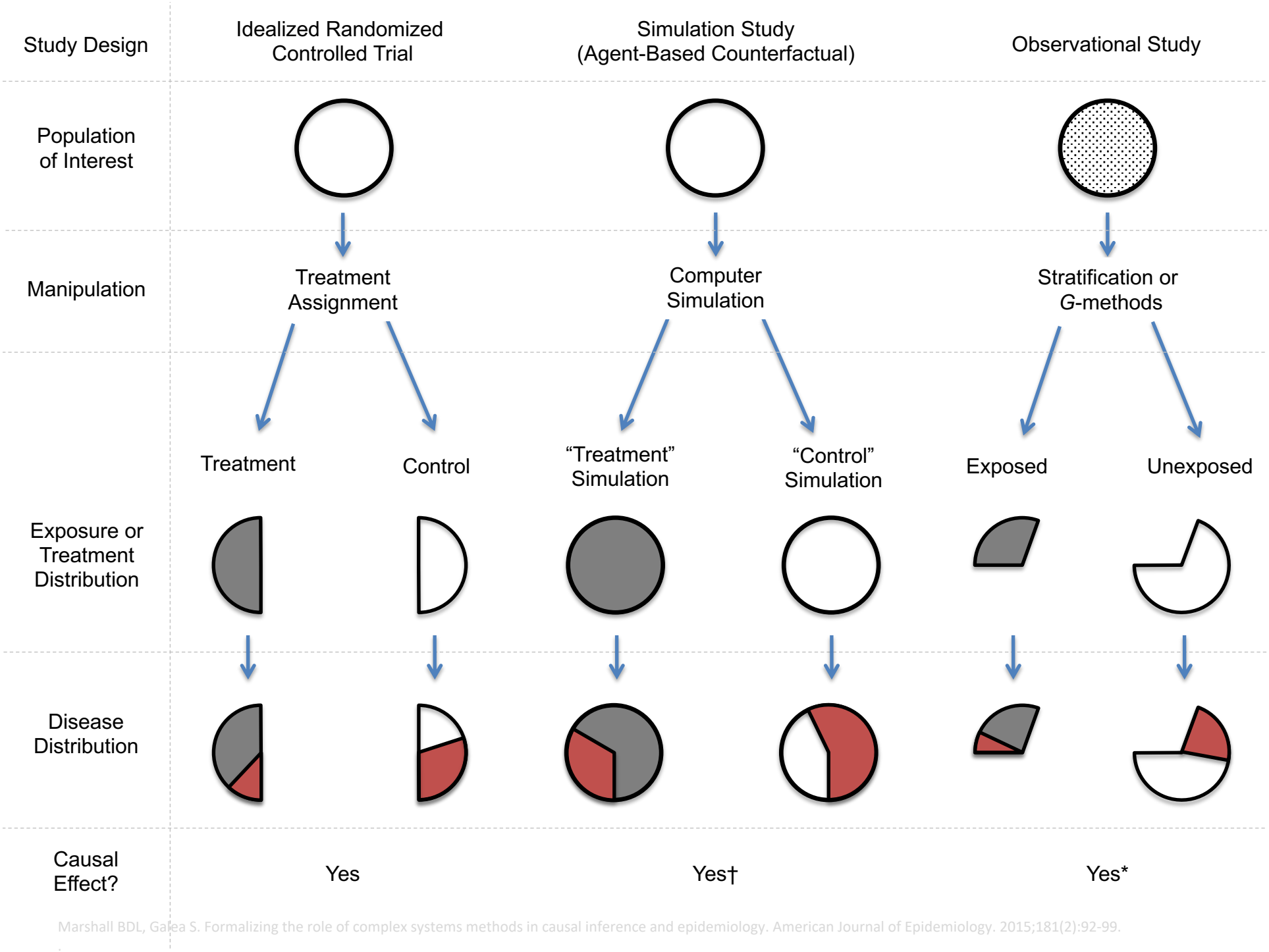




Marshall BDL, Galea S. Formalizing the role of complex systems methods in causal inference and epidemiology. American Journal of Epidemiology. 2015;181(2):92-99.



Marshall BDL, Galea S. Formalizing the role of complex systems methods in causal inference and epidemiology. American Journal of Epidemiology. 2015;181(2):92-99.



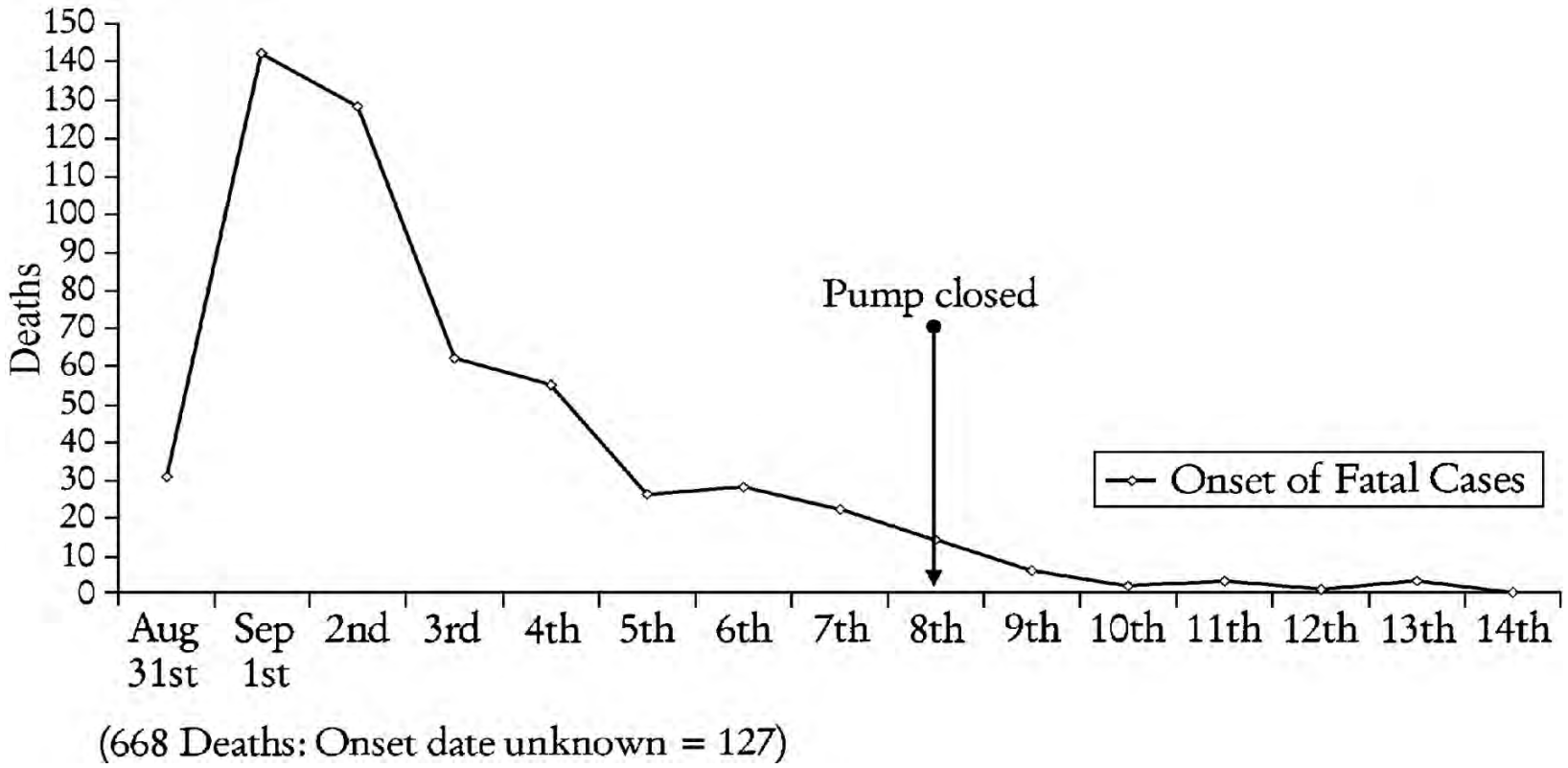
Marshall BDL, Galea S. Formalizing the role of complex systems methods in causal inference and epidemiology. American Journal of Epidemiology. 2015;181(2):92-99.

> Everything should be made as simple as possible, but not simpler

Attributed to Albert Einstein.

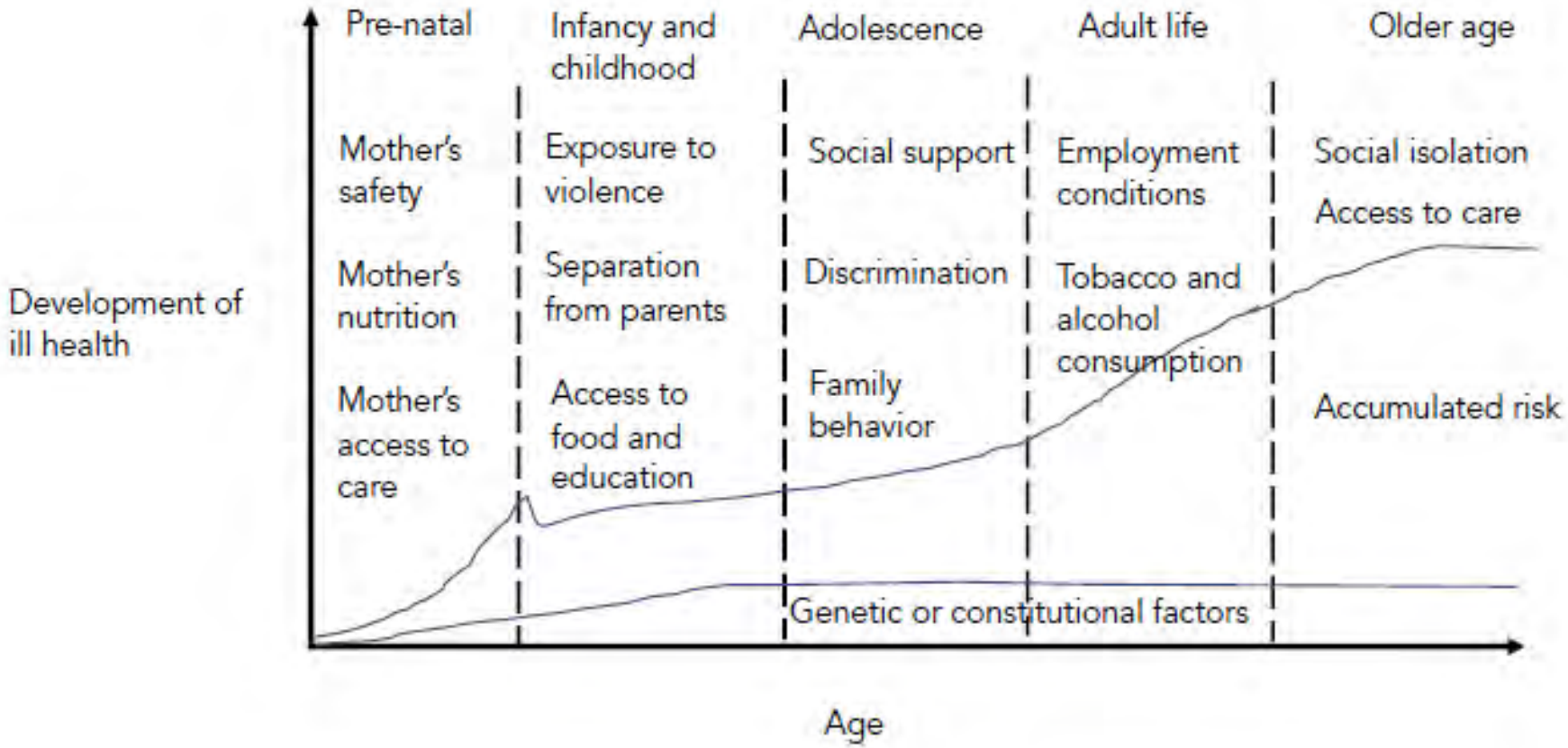
Simple approaches, a foundational myth

The effectiveness of simple approaches?



Davey Smith G, Ebrahim S. Epidemiology-is it time to call it a day? International Journal of Epidemiology. 2001;30:1-11.

5. The big picture


Pre-natal
Infancy and childhood
Adolescence
Adult life
Older age
Mother's safety
Mother's nutrition
Mother's access to care
Exposure to violence
Separation from parents
Access to food and education
Social support
Discrimination
Family behavior
Employment conditions
Tobacco and alcohol consumption
Social isolation
Access to care
Accumulated risk
Genetic or constitutional factors
Development of ill health
Age

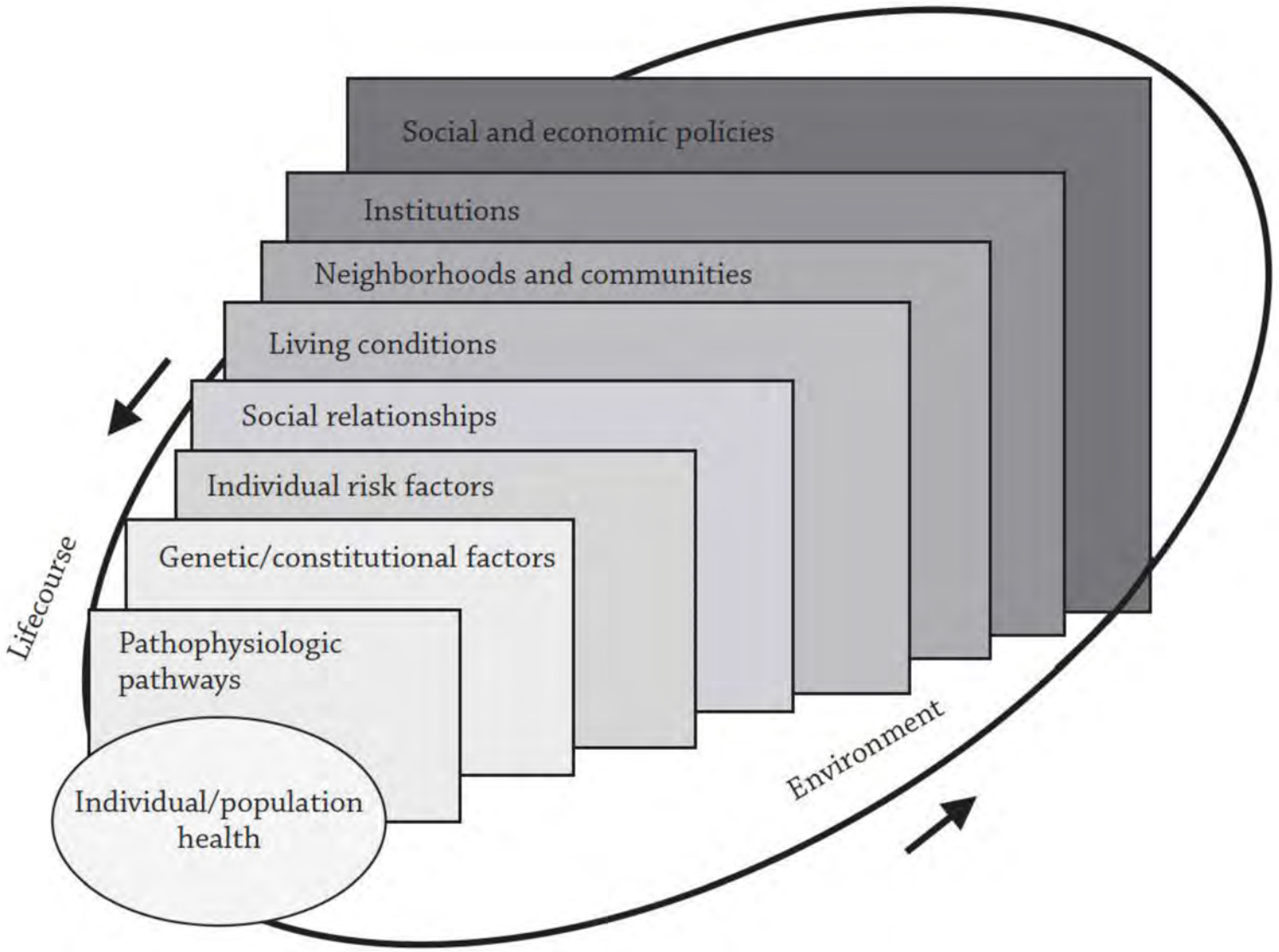
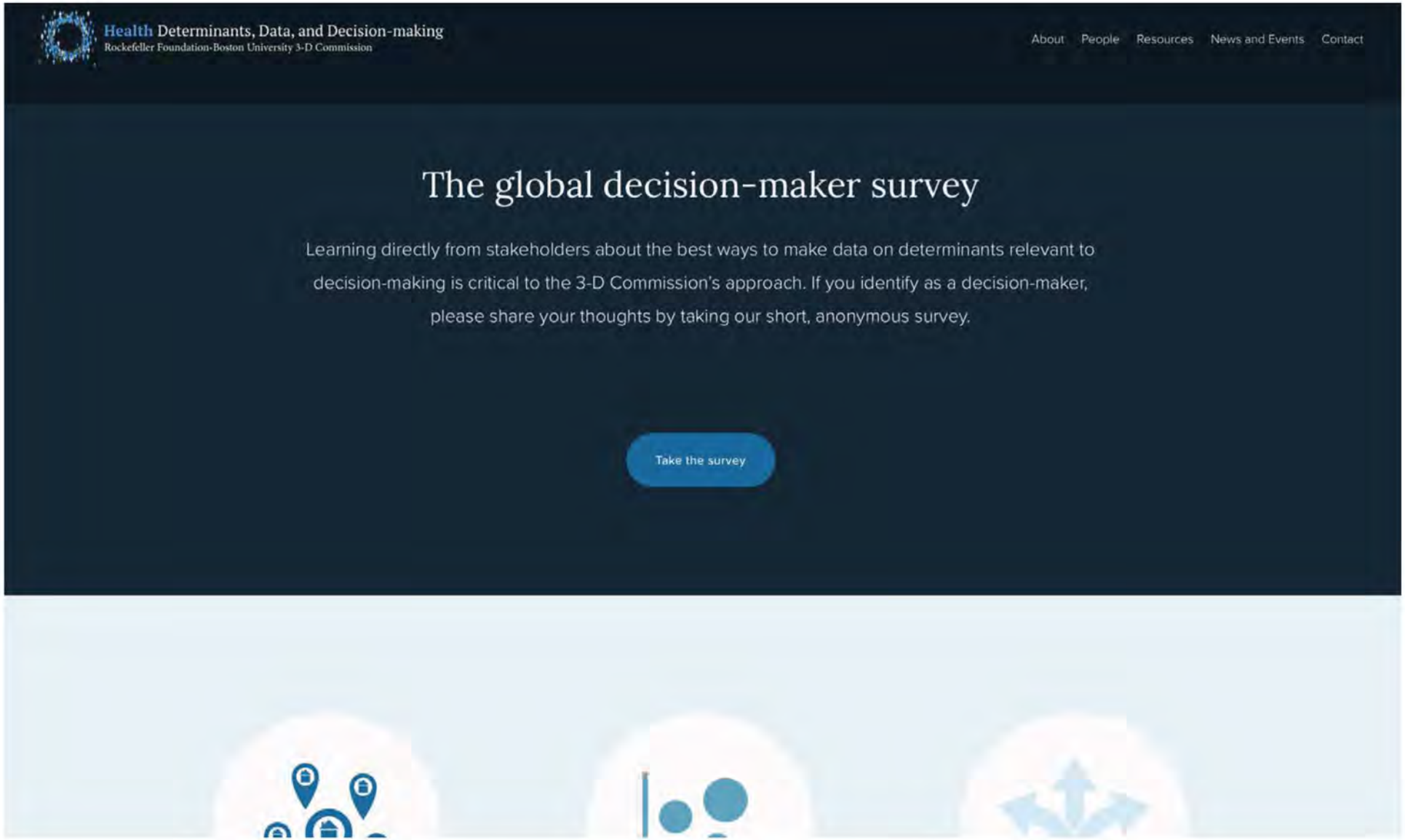


Figure 2.1 **Levels of influence on the health of populations.** Source: Modified from Kaplan, G. What's wrong with social epidemiology, and how can we make it better? *Epid Rev*. 2004; 26:124–135.


Health Determinants, Data, and Decision-making
Rockefeller Foundation-Boston University 3-D Commission
About
People
Resources
News and Events
Contact
The global decision-maker survey
Learning directly from stakeholders about the best ways to make data on determinants relevant to decision-making is critical to the 3-D Commission's approach. If you identify as a decision-maker, please share your thoughts by taking our short, anonymous survey.
Take the survey

twitter/@sandrogalea

sgalea@bu.edu