Understanding socio-structural drivers of HIV transmission using epidemiology and systems science

ANNA HOTTON, PHD, MPH

Research Assistant Professor
Department of Medicine
Chicago Center for HIV Elimination
University of Chicago, Chicago IL

THE UNIVERSITY OF CHICAGO

CCHE
The Chicago Center for HIV Elimination

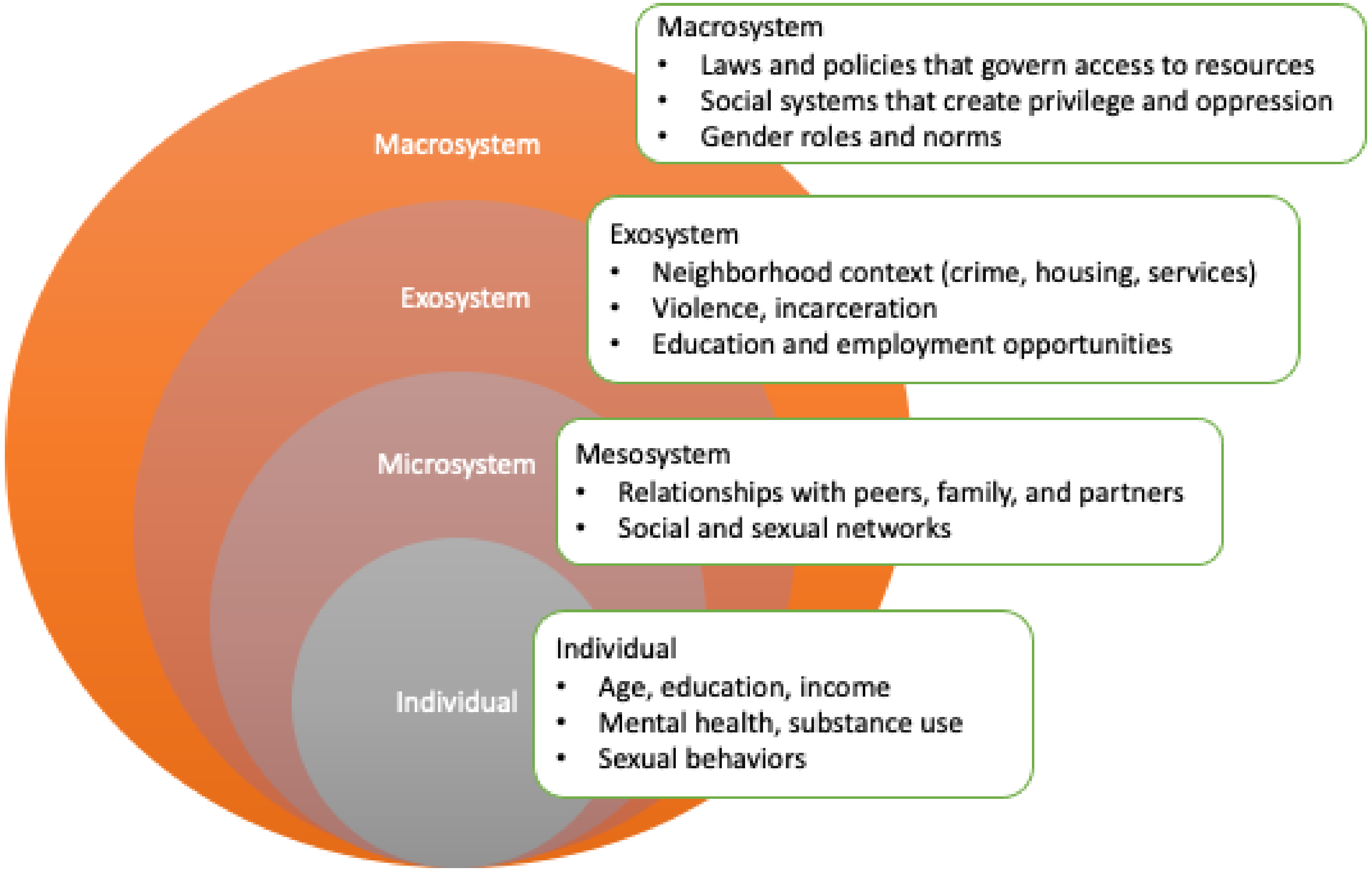
Outline

- Introduction
- HIV in Chicago
- Brief overview of agent-based modeling (ABM)
- ABM and counterfactual frameworks
- Example: ABM to understand the impact of criminal justice involvement on HIV transmission
- Future research and next steps

Introduction

- Socio-structural & contextual influences
  - Dyad-level factors
  - Network influences, environmental context
  - Social determinants of health (SDOH)

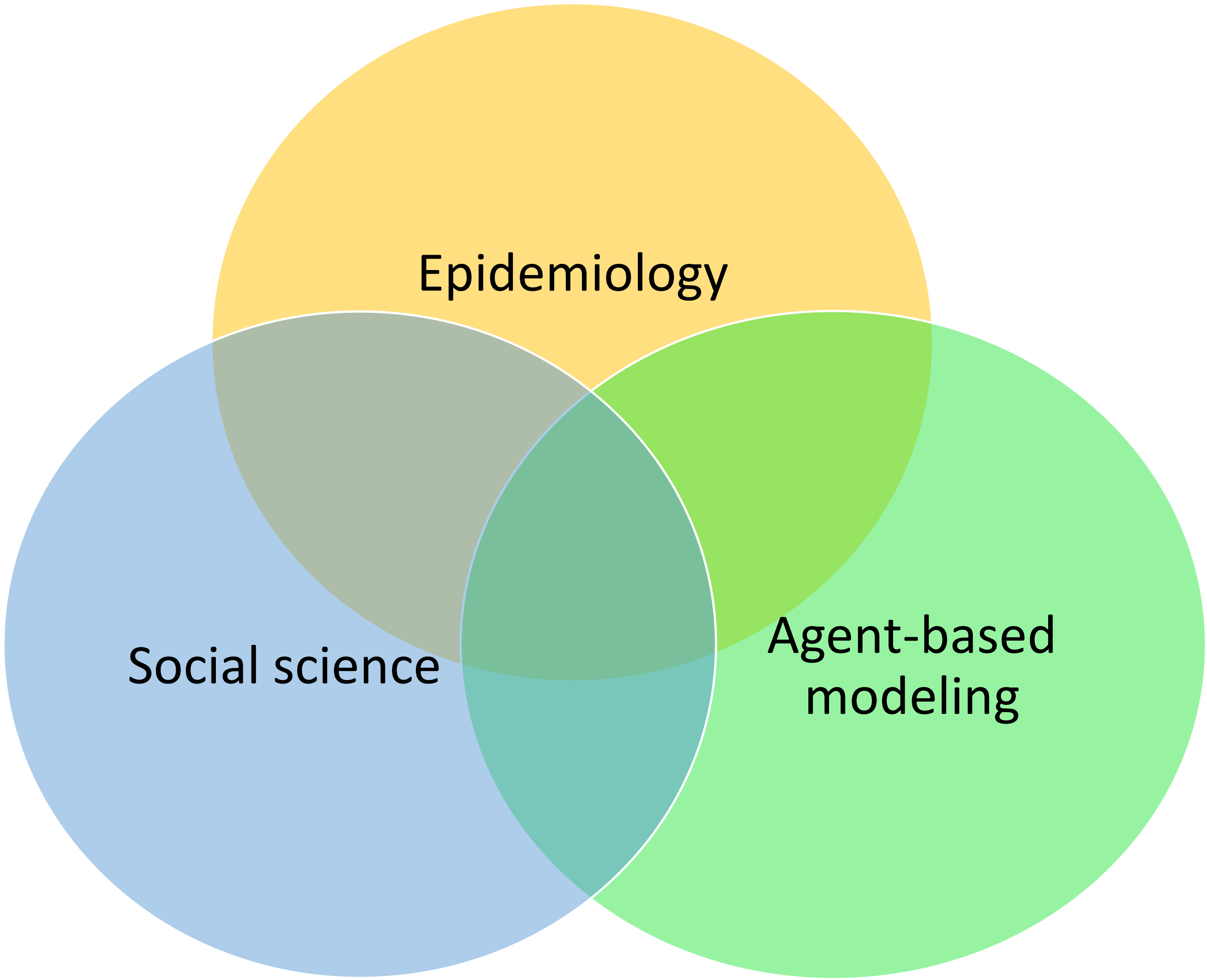
- Populations disproportionately affected by HIV/STIs



Adapted from Ecological Systems Theory. Source: Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press.

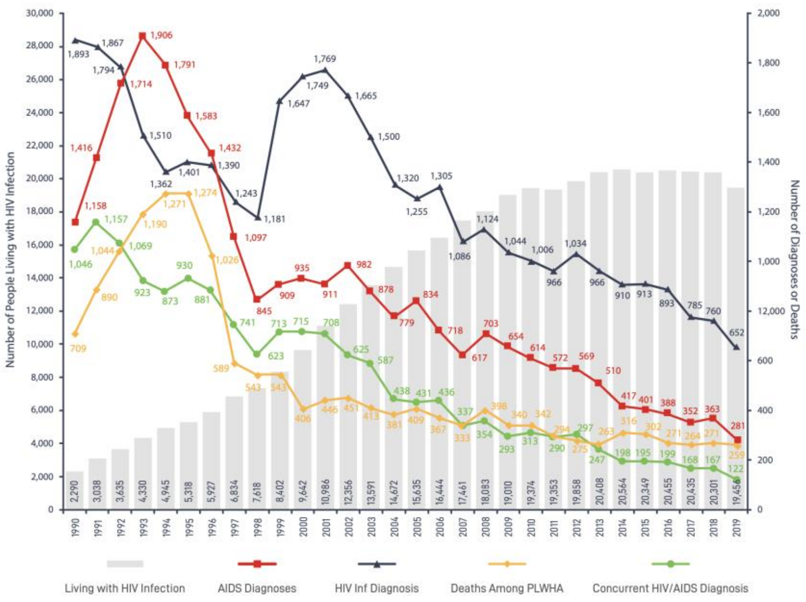
Research Methods

- Epidemiology
  - Observational research, interventions
  - Causal inference
- Social sciences
  - Conceptual frameworks
  - Social determinants of health
- Systems science & agent-based modeling
  - Useful for studying complex systems
  - Epidemiologic analyses provide input parameters for agent-based models



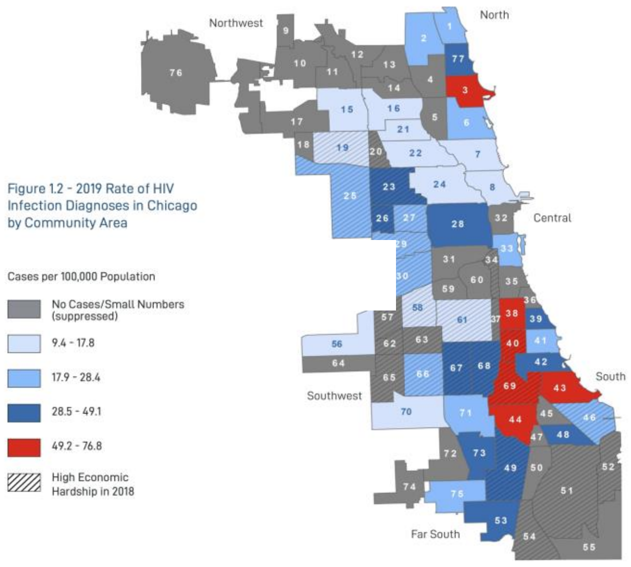
HIV in Chicago

Trends in HIV/AIDS Infections, Diagnoses, and Deaths, Chicago, 1990-2019

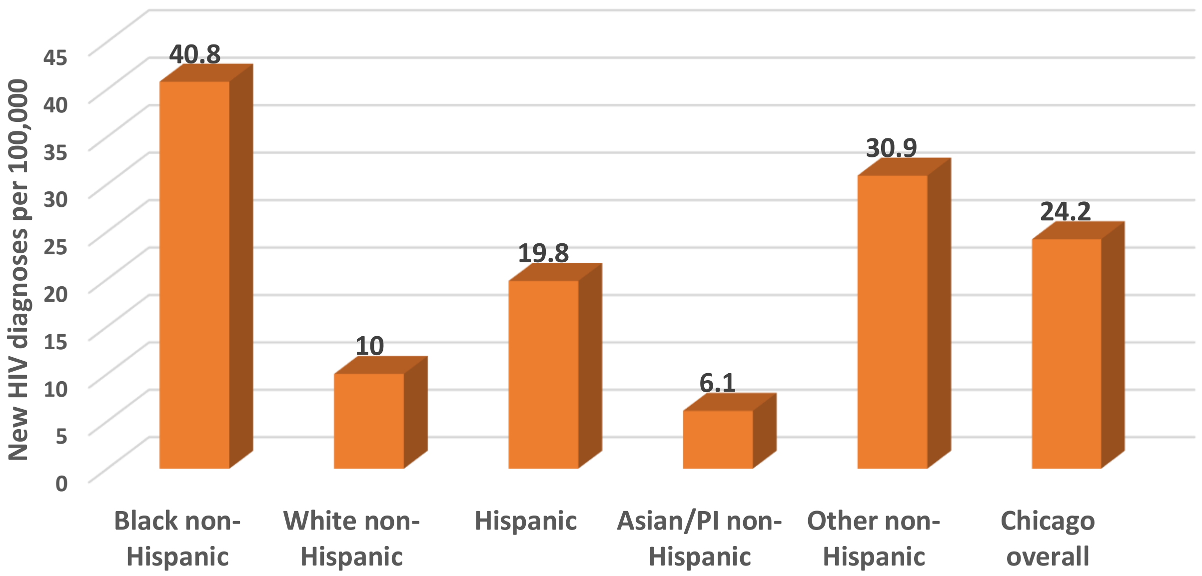


Source: Chicago Department of Public Health. HIV/STI Surveillance Report, 2019. Chicago, IL: City of Chicago; December 2020.

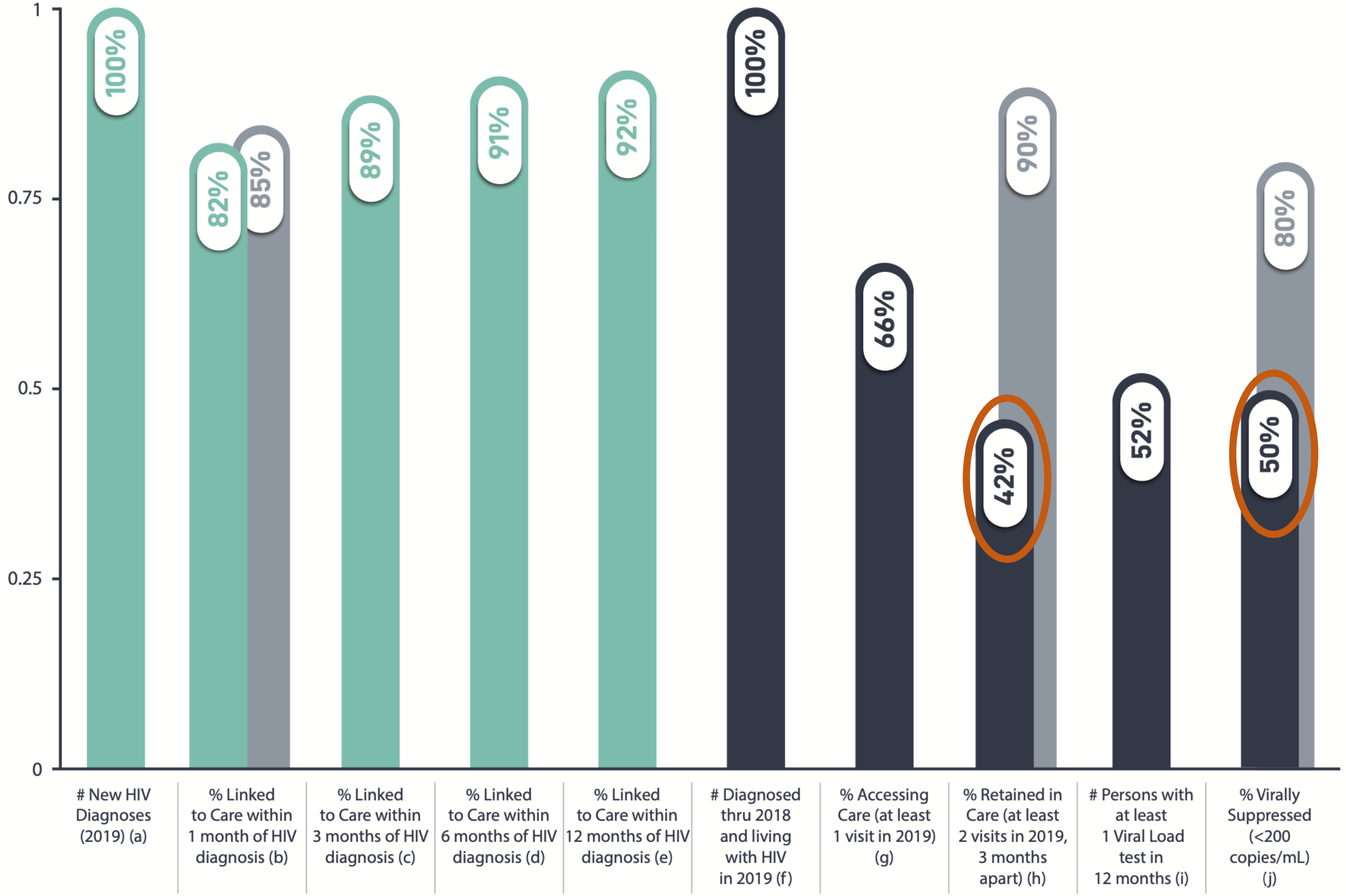
HIV Infection Rates by Chicago Community Area, 2019



New HIV diagnoses per 100,000, Chicago, 2019



HIV Continuum of Care Among Persons Aged ≥13 Years, Chicago, 2019



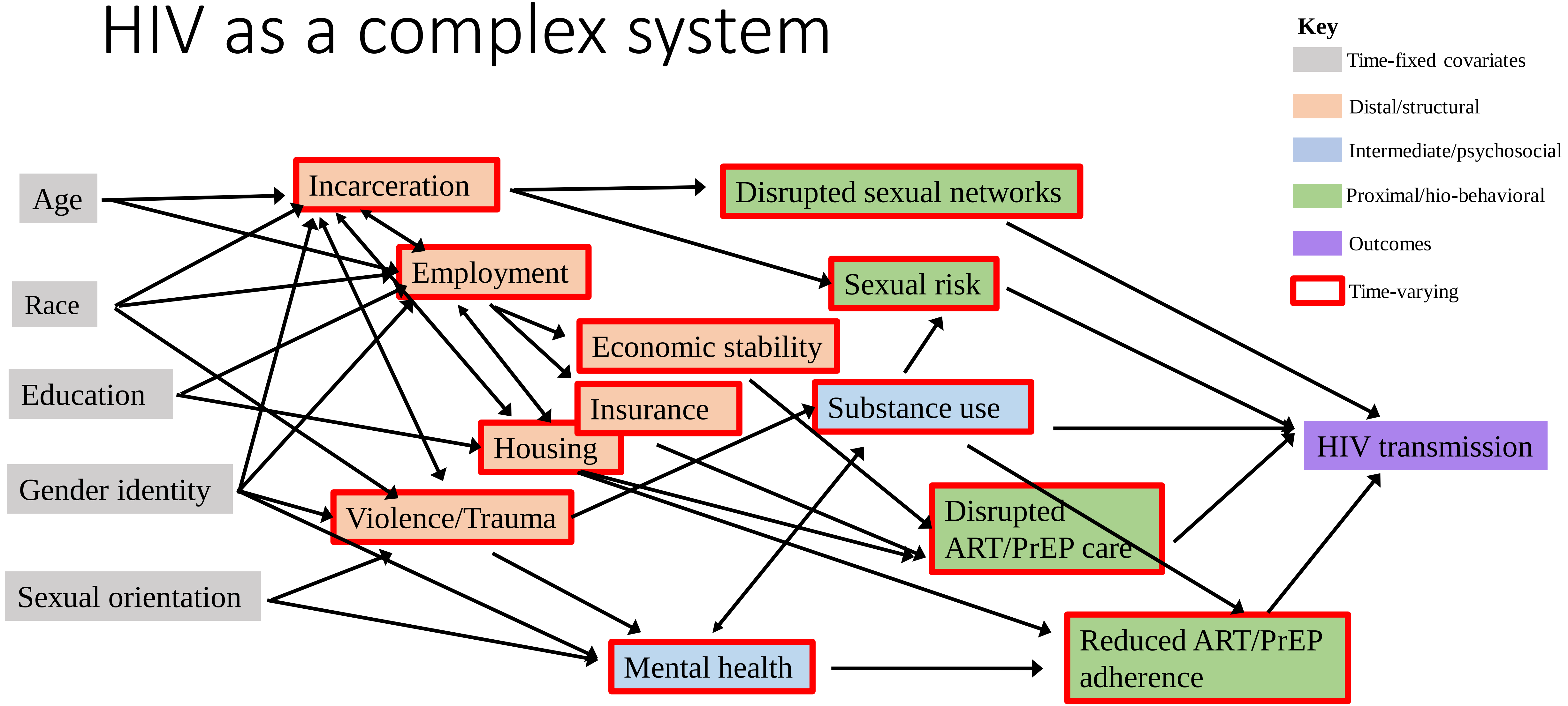
Source: Chicago Department of Public Health. HIV/STI Surveillance Report, 2019. Chicago, IL: City of Chicago; December 2020.

HIV Elimination goals

- Interventions will need to focus on communities with complex and co-occurring socio-structural barriers to engagement in HIV prevention and care

- Much previous research focused on behavior change at the individual level (sexual risk, substance use), but there is a recognized need to focus on more distal influences on HIV transmission
  - E.g., Housing, employment, incarceration

HIV as a complex system



Existing evidence and knowledge gaps

Evidence

- Associations of SDOH with HIV transmission related factors across multiple observational studies
  - Sexual behaviors
  - Substance use, mental health
  - Engagement in care, viral suppression, PrEP uptake

Gaps

- Many cross-sectional studies
- Focused on single factors or pathways
- Lacked statistical power or sufficient confounder control; causal associations cannot be inferred
- Effect size magnitude varies widely due to differences in study populations, design, timeframe, confounder control

Limitations of traditional study designs for understanding complex systems

- Logistical
  - Long duration of follow up required to observe effects
  - Very resource intensive to follow people longitudinally
  - Settings may not be conducive to traditional research designs (e.g., criminal justice settings)

- Ethical
  - Not always feasible or possible to randomize people
  - High participant burden/burnout

Statistical challenges

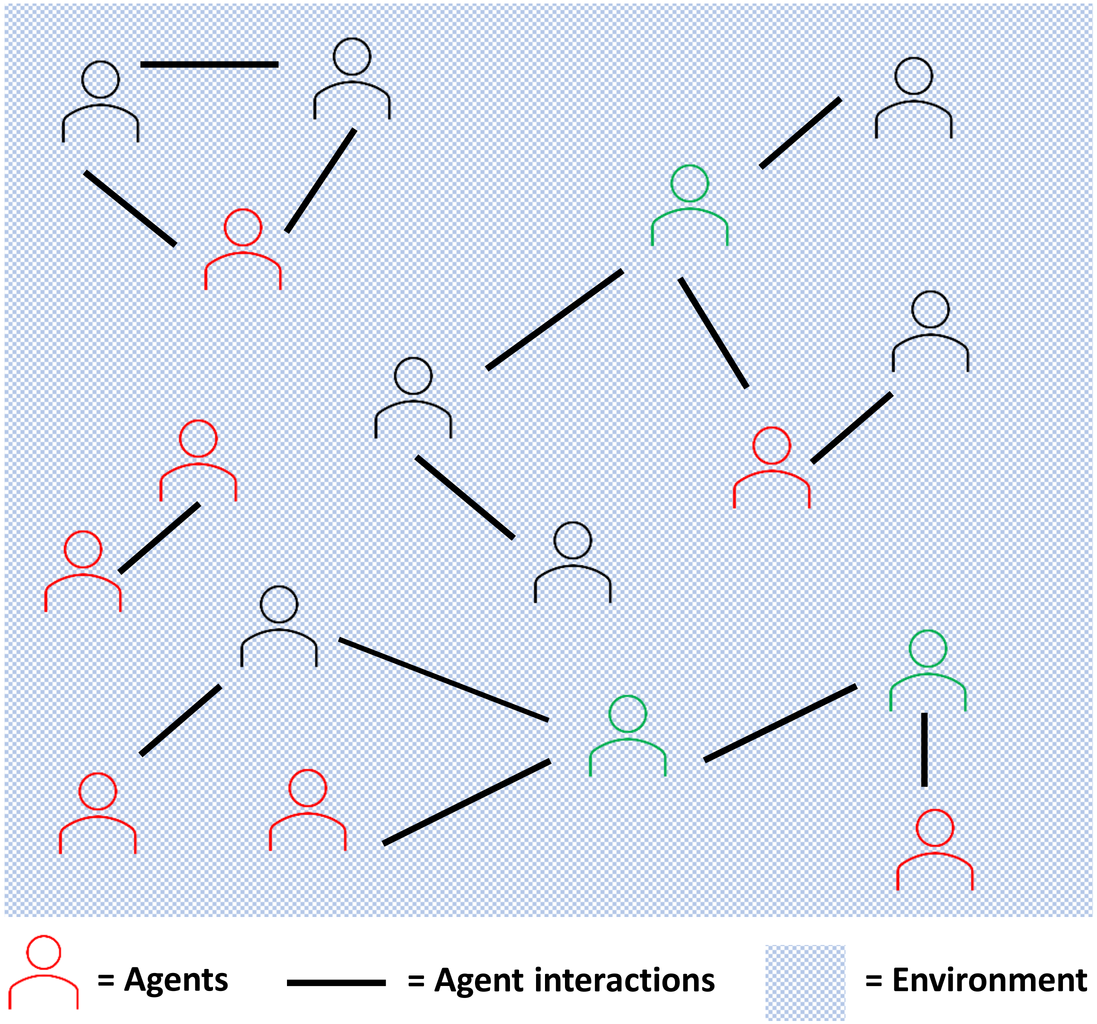
- Standard regression models typically assume relationships between exposures and outcomes are unidirectional, linear, and time-constant and exposures are independent
  - Not well suited to relationships characterized by causal interdependence, non-linearity (e.g., thresholds), feedback loops (magnified or dampening effects), and interference (one person's exposure influences the outcomes of others)
  - Many traditional causal inference frameworks assume these are absent (e.g., unidirectionality, no interference, etc.)
- For rare outcomes (e.g., HIV), very large sample sizes required for sufficient power

Limitations of randomized controlled trials (RCTs) for evaluating complex interventions

- RCTs – useful for isolating a single intervention effect or component, generally while holding other factors and contexts constant
- Systems science approaches are better suited to studying complex interventions
  - Questions that couldn’t be answered with simpler designs or models

Agent-Based Models (ABMs)

- Computer simulation approach to modeling the dynamics of complex systems
- Models represent social systems composed of agents that interact with and influence each other
- Observe system-level consequences of agent behaviors and interactions
- Effects of interventions can be simulated under various assumptions in a virtual environment



Agent-based model components

Agents	Agent-Agent interactions	Environments	Agent-environment interactions

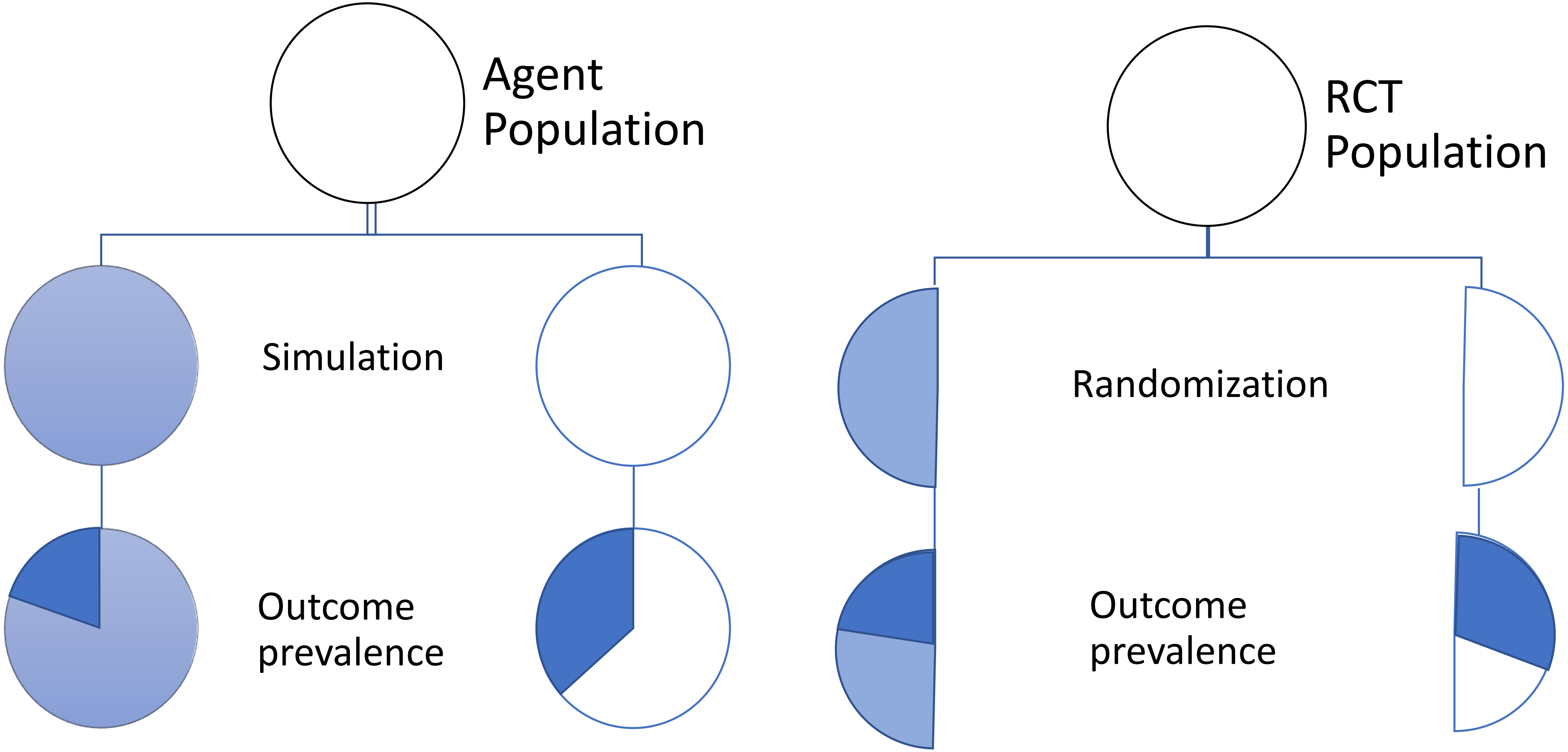
• Attributes (age, sex, race, employment, housing)
  • Static or dynamic
• Behaviors
  • Based on current information & past history
	• Information exchange
	• Disease transmission
	• Contend for resources
		• Social or sexual networks
		• Physical, social, neighborhood environments
			• Take in information on environment
			• Shape environment

Uses of agent-based modeling in epidemiology

- Understand mechanisms by which exposures (e.g., SDOH) impact population level health outcomes
  - Can show how patterns at the population level arise from exposures that might not be evident in a single study
  - Conduct counterfactual experiments to evaluate hypotheses that may not be possible with standard statistical models

- Evaluate potential interventions
  - Mechanisms by which interventions work
  - How interventions can be most efficiently focused (identify subgroups)
  - Optimal combination/sequence of interventions

ABM and counterfactual frameworks



Adapted from Marshall BD, Galea S. Formalizing the role of agent-based modeling in causal inference and epidemiology. Am J Epidemiol. 2015 Jan 15;181(2):92-9.

Basic notation: Agents

- At each time step t ($t = 1, \ldots, T$), each agent i ($i = 1, \ldots, N$) has a set of m ($m = 1, \ldots, M$) internal traits that can be described by the matrix $\mathbf{S}^{\mathrm{t}}$

$$\mathbf{S}^{\mathrm{t}} = \begin{bmatrix} s_{1,1}^{t} & s_{1,2}^{t} & \cdots & s_{1,M}^{t} \\ s_{2,1}^{t} & s_{2,2}^{t} & \cdots & s_{2,M}^{t} \\ \vdots & \vdots & \vdots & \vdots \\ s_{N,1}^{t} & s_{N,2}^{t} & \cdots & s_{N,M}^{t} \end{bmatrix}$$

- Traits:
  - Continuous, nominal, dichotomous
  - Can represent sociodemographics, exposures, behavioral proclivity, etc.

- Analogously, agents can be placed in p ($p = 1, \ldots, P$) environments where $\mathbf{E}^{\mathrm{t}}$ represents an environmental state matrix

Example from Marshall BD, Galea S. Formalizing the role of agent-based modeling in causal inference and epidemiology. Am J Epidemiol. 2015 Jan 15;181(2):92-9.

Agent-agent interactions

- At each time step t ($t = 1, …., T$) each agent i interacts with a subset of the population $\{1, …, i - 1, i + 1, …, N\}$
    - Described by agent-agent interaction matrix $\mathbf{K}^t$ where each element $k_{i,j}^t$ indicates whether agent *i* interacts with agent *j* during timestep *t* where i *and* $j = 1, …., N$
    - Can be symmetric or asymmetric (information or disease transmission can flow one way or bidirectionally)

Steps

- Initialize ABM by populating agent trait matrix, environmental state matrix, and interaction matrix with values from pre-defined probability distributions and functions

- Define set of rules $\mathbf{Z}$ for updating of agent traits, agent-agent interactions, and movement between or interaction with environments
  - Rules $\mathbf{Z}$ are defined by functions

- At each time step: update the model based on previous values and pre-defined rules $\mathbf{Z}$

Steps

- Monte Carlo simulations to obtain outcomes (e.g., disease incidence, prevalence, mortality) from runs r ($r = 1, \ldots, R$) at time T for counterfactual scenarios of interest:
  - Scenario A ($[\mathbf{Z}, \mathbf{S}_A^T, \mathbf{K}_A^T, \mathbf{E}_A^T]$) vs. Scenario B ($[\mathbf{Z}, \mathbf{S}_B^T, \mathbf{K}_B^T, \mathbf{E}_B^T]$)

- Compute point estimates from scenario A vs. scenario B by averaging across runs for each outcome of interest

$$\hat{\mu}_{o,A}^{T,R} = \frac{\sum_{r=1}^{R} y_{r,o,A}^{T}}{R} \text{ vs. } \hat{\mu}_{o,B}^{T,R} = \frac{\sum_{r=1}^{R} y_{r,o,B}^{T}}{R}$$

Example: Agent-based modeling to study the impact of criminal justice involvement (CJI) on HIV transmission among young Black sexual and gender minorities (SGM)

R01DA033934 (Fujimoto, Harawa, & Schneider, PIs)

R21MH128116 (Hotton, PI)

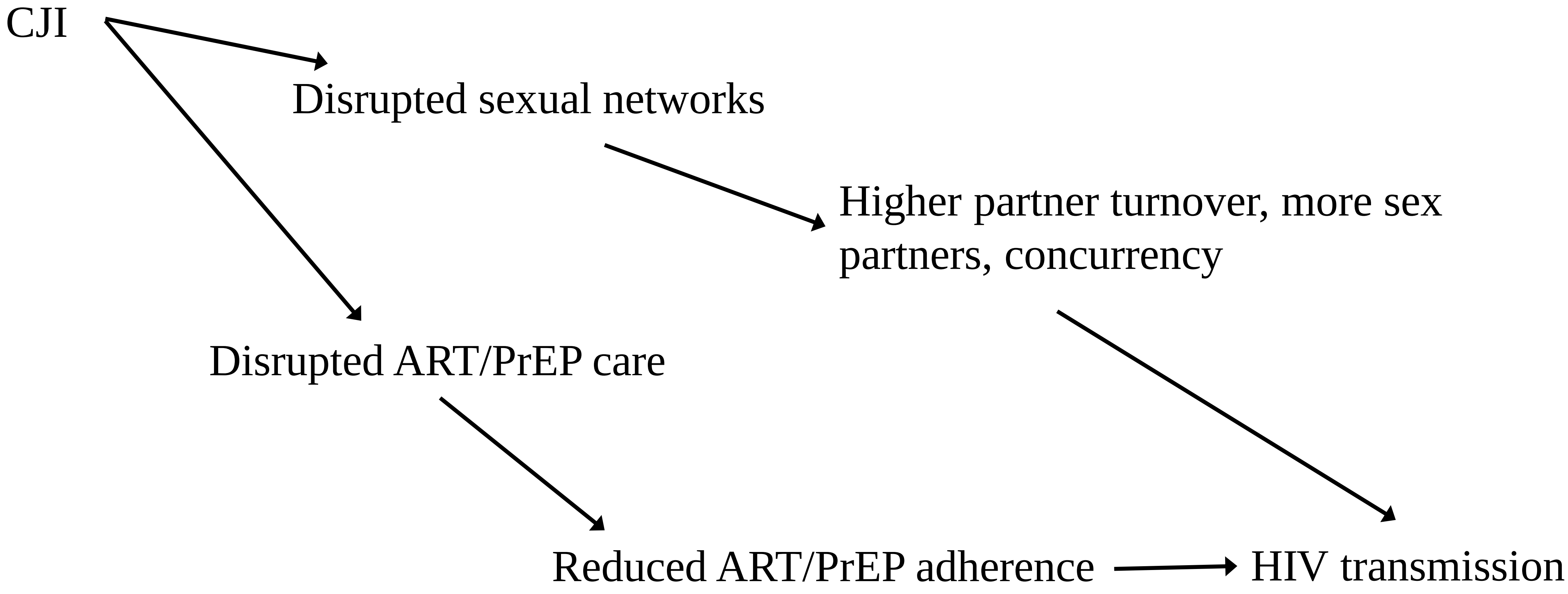
Intersection of HIV and criminal justice involvement

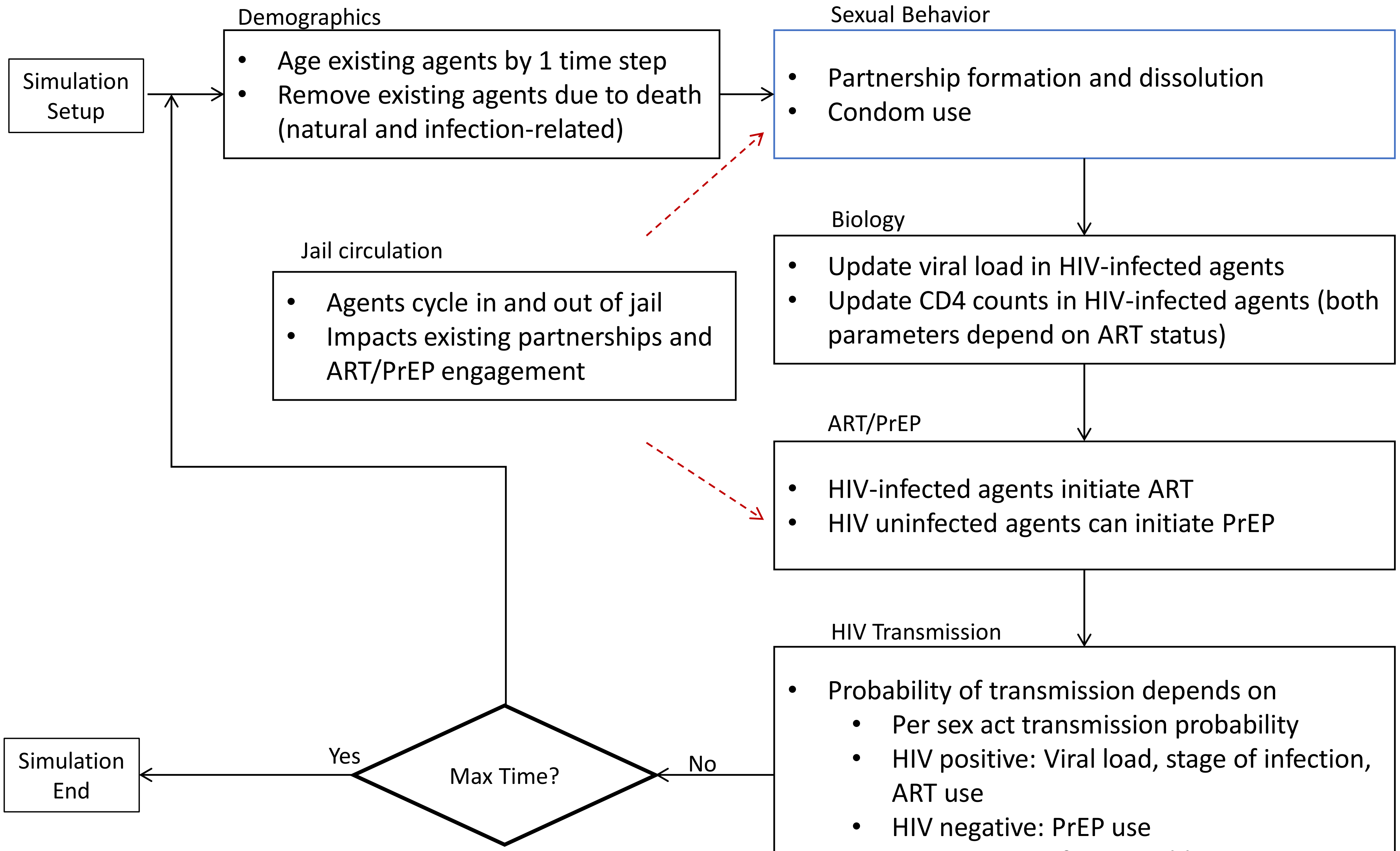
- Black SGM are disproportionately impacted by HIV and criminal justice involvement
  - Frequent cycling between communities and criminal justice settings

- CJI can impact:
  - Employment and housing opportunities
  - Access to medical care
  - Social and sexual network stability

- Agent-based models can be used to:
  - Provide insights to understand how CJI impacts HIV transmission
  - Evaluate interventions for criminal justice involved individuals and their networks

Hypothesized mechanisms by which justice involvement impacts HIV transmission




Simulation Setup
Demographics
Age existing agents by 1 time step
Remove existing agents due to death (natural and infection-related)
Sexual Behavior
Partnership formation and dissolution
Condom use
Jail circulation
Agents cycle in and out of jail
Impacts existing partnerships and ART/PrEP engagement
Biology
Update viral load in HIV-infected agents
Update CD4 counts in HIV-infected agents (both parameters depend on ART status)
ART/PrEP
HIV-infected agents initiate ART
HIV uninfected agents can initiate PrEP
HIV Transmission
Probability of transmission depends on
Per sex act transmission probability
HIV positive: Viral load, stage of infection, ART use
HIV negative: PrEP use
Circumcision of susceptible partners
Max Time?
Yes
No
Simulation End

Model population and data sources

- Model population: 10,000 agents representing Black SGM ages 18-34 in the city of Chicago
- Data sources: Local cohort studies, clinical data, and public health surveillance
- Outcomes: HIV incidence and prevalence (average annualized estimates computed over 10 years)
- Calibration: local HIV incidence and prevalence estimates (CDPH surveillance and local studies), incarceration incidence and prevalence
- Model components: Developed with Repast HPC ABM toolkit using C++
- Network formation and dissolution dynamics modeled with exponential random graph models using the *statnet* suite of packages in R

Khanna et al. A modeling framework to inform preexposure prophylaxis initiation and retention scale-up in the context of 'Getting to Zero' initiatives. AIDS 2019, 33(12): 1911-1922.

Experiments

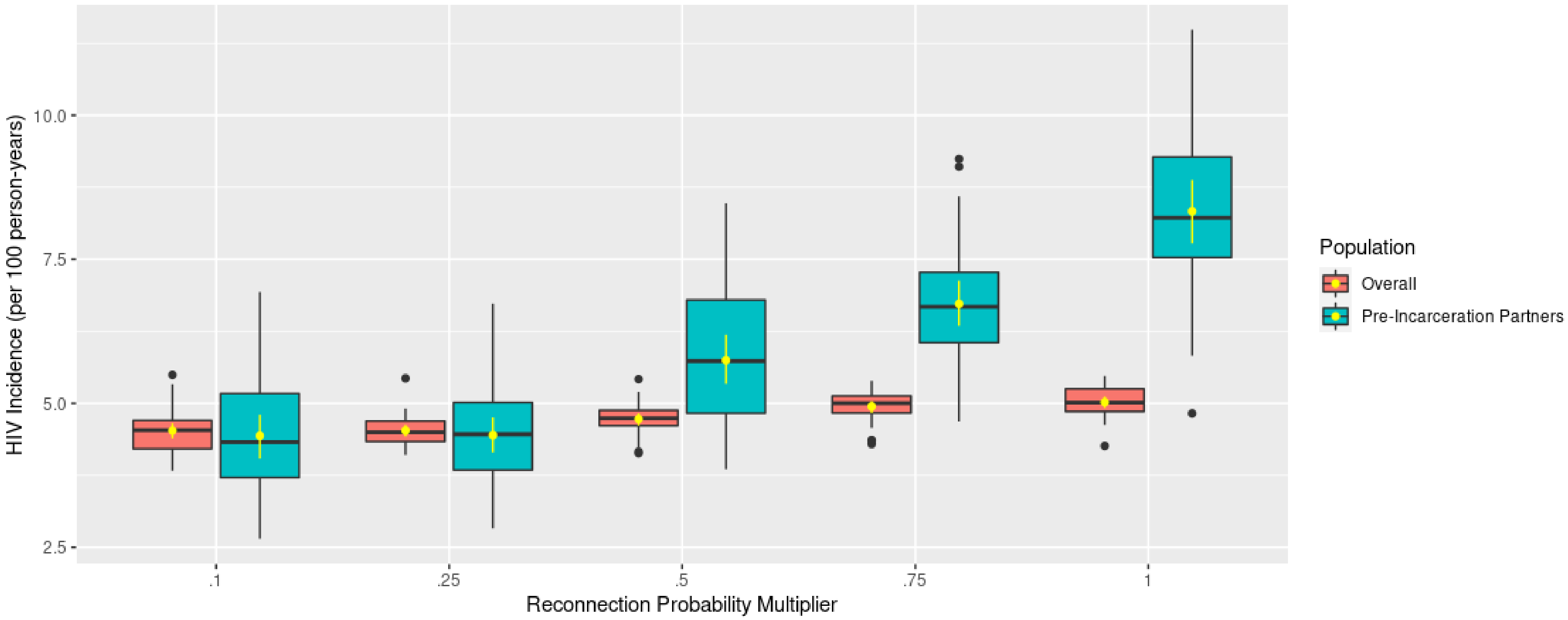
- Quantify the impact of criminal justice involvement:
  - Population level HIV incidence
  - HIV incidence among justice involved individuals and their networks
- Evaluate interventions to reduce the impact of justice involvement on HIV transmission in different sub-populations
  - Examples:
    - Reduce post-release disruption in HIV/PrEP care (e.g., interventions to facilitate care engagement by reducing insurance, housing, or employment barriers)
    - Focused or enhanced PrEP and ART interventions for justice involved individuals and their networks
- Each scenario was repeated across 30 runs to incorporate stochasticity and outcomes were averaged across runs

Annual HIV incidence by subpopulation

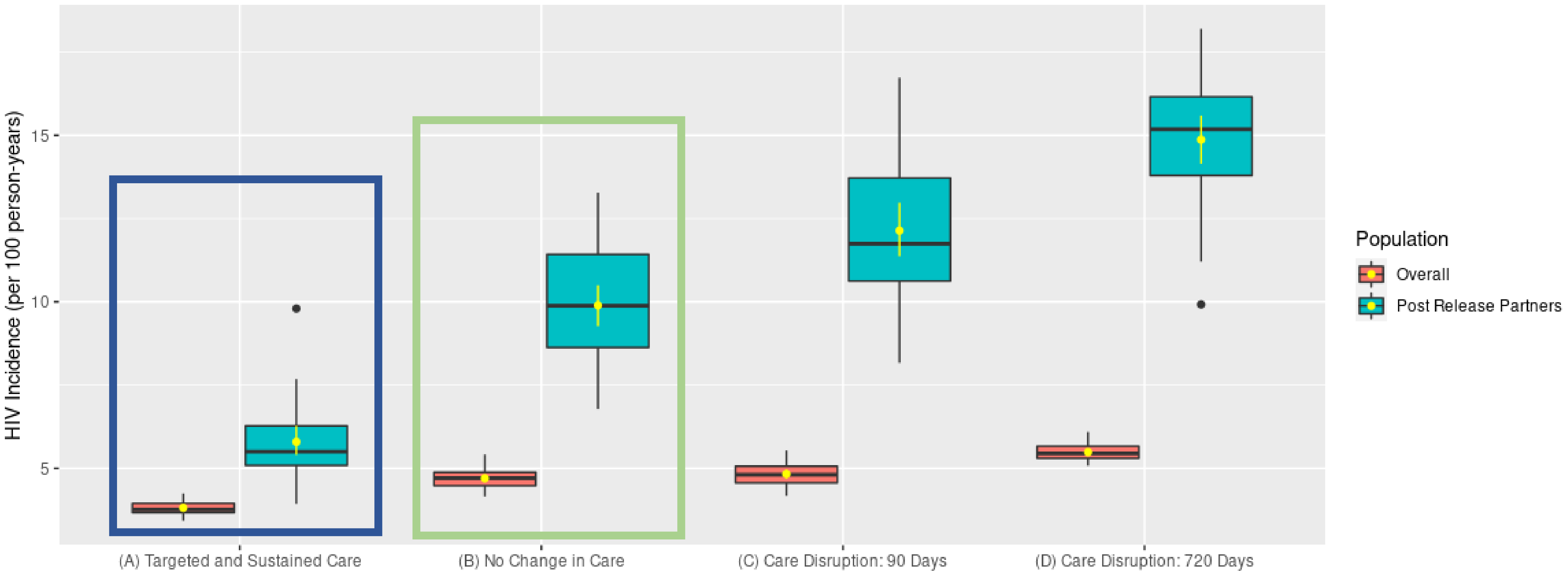
Population	**HIV incidence per 100 person-years**	**95% CI***
Individuals		
Ever incarcerated	5.72	5.52 – 5.92
Never incarcerated	4.73	4.65 – 4.83
Partners of individuals with CJI		
Pre-Incarceration Partners	6.83	6.31 – 7.40
Post-Release Partners	12.14	11.4 – 13.0

*Confidence intervals obtained via bootstrapping

HIV incidence by probability of partner reconnection after release from jail



HIV incidence under various levels of post-release care disruption and intervention



Average HIV incidence under different care disruption scenarios

	Post-Release Partners Incidence per 100 person-years (95% CI)	Overall Incidence per 100 person-years (95% CI)
Intervention: Targeted and sustained care	5.80 (5.40, 6.28)	3.81 (3.74, 3.89)
Standard: No change in care	9.90 (9.27, 10.50)	4.70 (4.59, 4.82)
Incidence rate ratio	0.59 (0.53, 0.65)	0.86 (0.84-0.89)
Incidence rate difference	-4.10 (-4.85, -3.35)	-0.89 (-1.02, -0.76)

Summary

- Identified a subgroup who could benefit from targeted PrEP interventions (partners of those recently released from jail) which may have been hard to identify using observational research designs
  - Can give ideas about where to target limited public health resources

- Suggests need for interventions to increase ART and viral suppression among HIV-positive individuals with CJI *and* increase PrEP/ART use in their networks

- Next steps
  - Evaluate impact of interventions to distribute PrEP to networks of released individuals
  - Explicitly model interventions to reduce post-release disruption in care by reducing insurance, housing, or employment barriers and recidivism
  - Combinations of interventions applied simultaneously or sequentially

Limitations & open questions

- Estimates used as input parameters for agent-based models are often uncertain or potentially biased – need for sensitivity analysis
  - Model results may be dependent on parameter inputs for which the true magnitude of effect is often unknown
  - Transportability – estimates of effect from one population may not generalize to another

- Can provide a range of effect estimates as priors and use computational techniques to refine estimates – ongoing work in this area
  - Large-scale sensitivity analyses and model exploration with high-performance computing
  - Identify variables that have the most impact on model (system) behavior

Extensions and ongoing work

- Ongoing work: extend existing model to incorporate additional social determinants of health (housing, employment) and evaluate their impact
  - Increase the granularity of the synthetic population in order to represent HIV transmission with sufficient realism to examine more nuanced research questions
- Develop formal methods for evaluating assumptions needed for valid inference with agent-based modeling
  - Counterfactual frameworks and high-dimensional sensitivity analyses to assess the impact of varying causal or mechanistic assumptions
  - Quantify the impact of incomplete or imprecise empirical data

- G-computation
  - Methods to estimate the causal effect of a time-varying exposure in the presence of time-varying confounders affected by the exposure; also applies to settings with feedback loops
  - Extensions can address interference (auto g-computation)
  - May complement agent-based modeling to triangulate information if adequate longitudinal data are available
  - Can provide causal effects estimates as starting parameters for agent-based models to be refined using computational approaches within the ABM

Collaborators and funding

- Aditya Khanna, Francis Lee, Daniel Sheeler, Nick Collier, Jonathan Ozik, Natascha Del Vecchio, Bryan Brickman, Russell Brewer, Kayo Fujimoto, Nina Harawa, John Schneider
- Funding: R01DA039934, R01DA033875, U2CDA050098, R21MH128116

Contact info

Anna Hotton, PhD, MPH

Research Assistant Professor, Department of Medicine

Director of Epidemiology, Chicago Center for HIV Elimination (CCHE)

University of Chicago

5837 S. Maryland Ave, L-038

Chicago, IL 60637

ahotton@medicine.bsd.uchicago.edu

Citations: Agent-based modeling, G-computation, and causal inference

Breger TL, Edwards JK, Cole SR, Westreich D, Pence BW, Adimora AA. Two-stage g-computation: Evaluating Treatment and Intervention Impacts in Observational Cohorts When Exposure Information Is Partly Missing. Epidemiology. 2020 Sep;31(5):695-703.

Breskin A, Edmonds A, Cole SR, Westreich D, Cocohoba J, Cohen MH, Kassaye SG, Metsch LR, Sharma A, Williams MS, Adimora AA. G-computation for policy-relevant effects of interventions on time-to-event outcomes. Int J Epidemiol. 2021 Jan 23;49(6):2021-2029.

Buchanan AL, Bessey S, Goedel WC, King M, Murray EJ, Friedman SR, Halloran ME, Marshall BDL. Disseminated Effects in Agent-Based Models: A Potential Outcomes Framework and Application to Inform Preexposure Prophylaxis Coverage Levels for HIV Prevention. Am J Epidemiol. 2021 May 4;190(5):939-948.

Buchanan AL, Hudgens MG, Cole SR, Mollan KR, Sax PE, Daar ES, Adimora AA, Eron JJ, Mugavero MJ. Generalizing Evidence from Randomized Trials using Inverse Probability of Sampling Weights. J R Stat Soc Ser A Stat Soc. 2018 Oct;181(4):1193-1209.

Buchanan AL, Vermund SH, Friedman SR, Spiegelman D. Assessing Individual and Disseminated Effects in Network-Randomized Studies. Am J Epidemiol. 2018 Nov 1;187(11):2449-2459.

Chatton A, Le Borgne F, Leyrat C, Gillaizeau F, Rousseau C, Barbin L, Laplaud D, Léger M, Giraudeau B, Foucher Y. G-computation, propensity score-based methods, and targeted maximum likelihood estimator for causal inference with different covariates sets: a comparative simulation study. Sci Rep. 2020 Jun 8;10(1):9219.

Crawford FW, Morozova O, Buchanan AL, Spiegelman D. Interpretation of the Individual Effect Under Treatment Spillover. Am J Epidemiol. 2019 Aug 1;188(8):1407-1409.

Dahabreh IJ, Haneuse SJA, Robins JM, Robertson SE, Buchanan AL, Stuart EA, Hernán MA. Study Designs for Extending Causal Inferences From a Randomized Trial to a Target Population. Am J Epidemiol. 2021 Aug 1;190(8):1632-1642.

Hernán MA. Invited commentary: Agent-based models for causal inference—reweighting data and theory in epidemiology. Am J Epidemiol. 2015 Jan 15;181(2):103-5.

Le Borgne F, Chatton A, Léger M, Lenain R, Foucher Y. G-computation and machine learning for estimating the causal effects of binary exposure statuses on binary outcomes. Sci Rep. 2021 Jan 14;11(1):1435.

Citations: Agent-based modeling, G-computation, and causal inference continued

Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing Study Results: A Potential Outcomes Perspective. Epidemiology. 2017 Jul;28(4):553-561.

Marshall BD, Galea S. Formalizing the role of agent-based modeling in causal inference and epidemiology. Am J Epidemiol. 2015 Jan 15;181(2):92-9.

Mooney SJ, Shev AB, Keyes KM, Tracy M, Cerdá M. G-computation and agent-based modeling for social epidemiology: Can population interventions prevent post-traumatic stress disorder?. Am J Epidemiol. 2021 Aug 18:kwab219. doi: 10.1093/aje/kwab219. Online ahead of print.

Murray EJ, Marshall BDL, Buchanan AL. Emulating Target Trials to Improve Causal Inference From Agent-Based Models. Am J Epidemiol. 2021 Aug 1;190(8):1652-1658.

Murray EJ, Robins JM, Seage GR 3rd, Freedberg KA, Hernán MA. The Challenges of Parameterizing Direct Effects in Individual-Level Simulation Models. Med Decis Making. 2020 Jan;40(1):106-111.

Murray EJ, Robins JM, Seage GR 3rd, Lodi S, Hyle EP, Reddy KP, Freedberg KA, Hernán MA. Using Observational Data to Calibrate Simulation Models. Med Decis Making. 2018 Feb;38(2):212-224.

Murray EJ, Robins JM, Seage GR, Freedberg KA, Hernán MA. A Comparison of Agent-Based Models and the Parametric G-Formula for Causal Inference. Am J Epidemiol. 2017 Jul 15;186(2):131-142.

Smith MJ, Mansournia MA, Maringe C, Zivich PN, Cole SR, Leyrat C, Belot A, Rachet B, Luque-Fernandez MA. Introduction to computational causal inference using reproducible Stata, R, and Python code: A tutorial. Stat Med. 2022 Jan 30;41(2):407-432.

Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. Am J Epidemiol. 2011 Apr 1;173(7):731-8.

Tchetgen Tchetgen EJ, Fulcher IR, Shpitser I. Auto-G-Computation of Causal Effects on a Network. J Am Stat Assoc. 2021;116(534):833-844. doi: 10.1080/01621459.2020.1811098.

Tennant PWG, Murray EJ, Arnold KF, et. al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021 May 17;50(2):620-632.

Tracy M, Cerdá M, Keyes KM. Agent-Based Modeling in Public Health: Current Applications and Future Directions. Annu Rev Public Health. 2018 Apr 1;39:77-94.

Wang A, Arah OA. G-computation demonstration in causal mediation analysis. Eur J Epidemiol. 2015 Oct;30(10):1119-27. doi: 10.1007/s10654-015-0100-z.

Wang A, Nianogo RA, Arah OA. G-computation of average treatment effects on the treated and the untreated. BMC Med Res Methodol. 2017 Jan 9;17(1):3. doi: 10.1186/s12874-016-0282-4.

Technical resources and documentation

Repast for High performance computing:

Homepage: https://repast.github.io/repast_hpc.html

Repast documentation, including Repast4Py (Python implementation) and lots of tutorials: https://repast.github.io/docs.html

Collier N, North M. Parallel agent-based simulation with Repast for High Performance Computing. *SIMULATION.* 2013;89(10):1215-1235.

Estreme-scale Model Exploration with Swift: https://emews.github.io/

Ozik J, Collier NT, Wozniak JM, Macal CM, An G. Extreme-Scale Dynamic Exploration of a Distributed Agent-Based Model With the EMEWS Framework. *IEEE Transactions on Computational Social Systems.* 2018;5(3):884-895.

Ozik J, Collier NT, Wozniak JM, Spagnuolo C. From desktop to Large-Scale Model Exploration with Swift/T. Paper presented at: 2016 Winter Simulation Conference (WSC); 11-14 Dec. 2016, 2016.

EpiModel (R package for mathematical models of infectious disease): https://www.epimodel.org/

Tutorials on agent-based modeling:

C M Macal (2016) Everything you need to know about agent-based modelling and simulation, Journal of Simulation, 10:2, 144-156, DOI: 10.1057/jos.2016.7

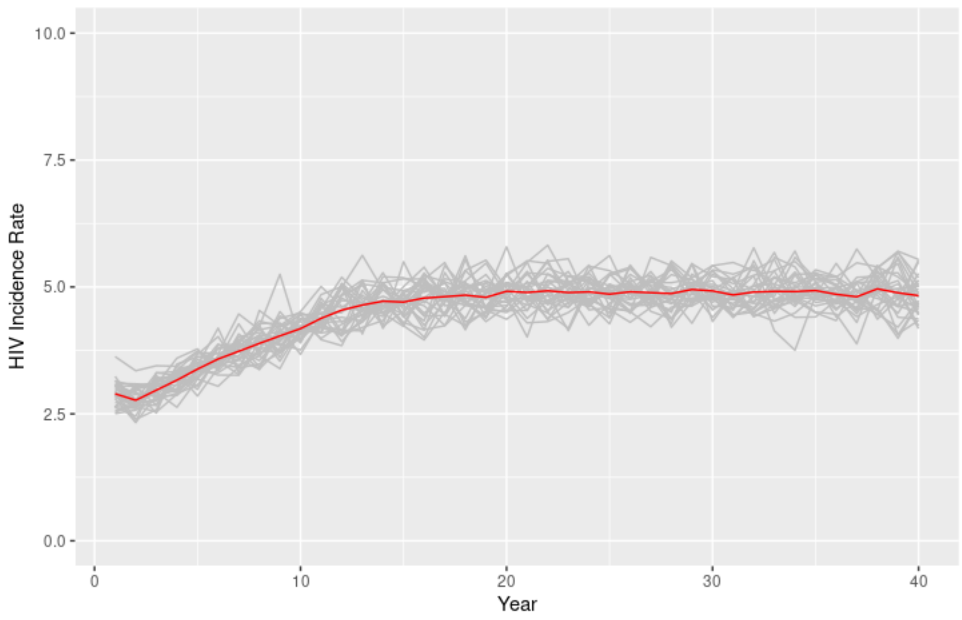
Macal & North: Introductory tutorial on agent-based modeling and simulation: http://simulation.su/uploads/files/default/2014-macal-north.pdf

Extra slides

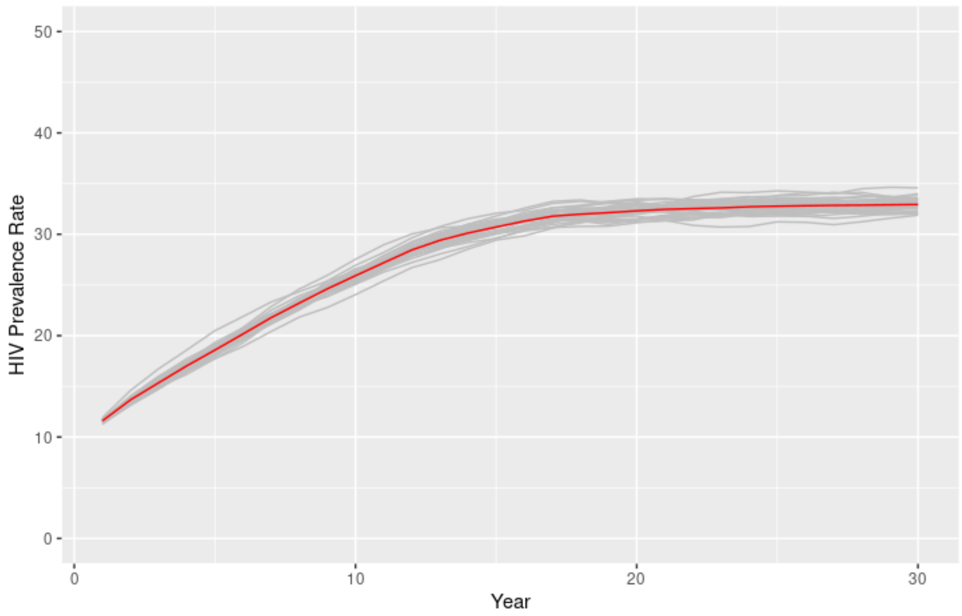
Model calibration

- Initial set of 270 calibration runs
- Calibration targets: annual HIV incidence (5-7%), HIV & incarceration prevalence
  - Examined differences by age and prior incarceration
- Tested a range of scenarios using empirical estimates from local data as inputs
  - Probability & duration of incarceration
  - PrEP & ART care continuum disruption
  - Network tie retention probabilities
- Refined estimates after initial examination of model output
- Selected the set of parameters that produced outputs consistent with empirical calibration targets for further experimentation

HIV incidence

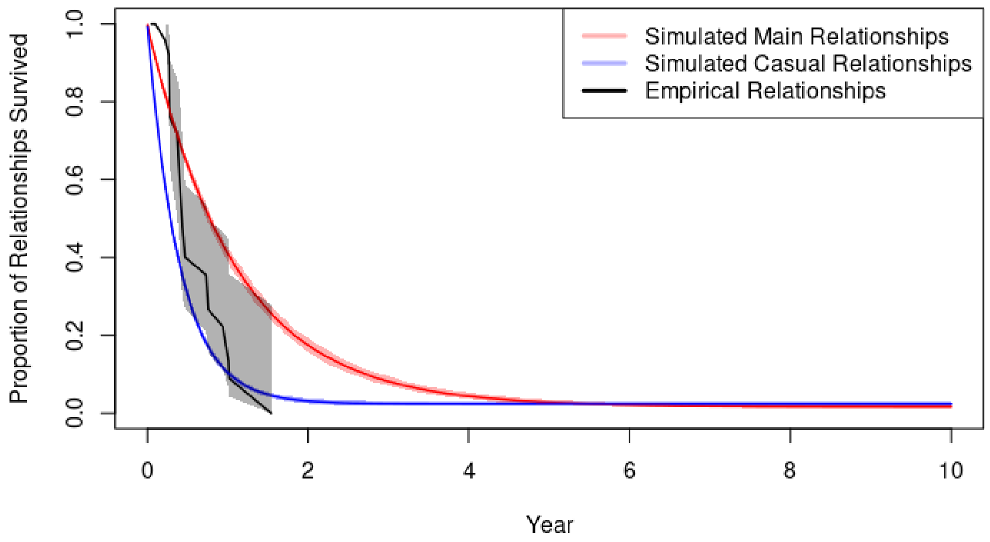


HIV prevalence



Network tie retention

- When agents go to jail existing network ties are broken with a probability of reconnecting after release
- Determine the status quo survival rate of relationship ties over time using the existing ABM with no incarceration processes implemented
- Apply multiplier to represent the impact of incarceration on probability of reconnecting ties after release
    - Results in a shift in the status quo distribution
        - Multiplier = 1: no impact on tie retention
        - Multiplier < 1: probabilities of retained ties less than the status quo



Local data sources

Source	Year	Description	Parameter categories
UConnect (R01DA033875)	2013-2015	Cohort study of Black MSM & transwomen ages 16-29, RDS recruitment (n=618)	Sociodemographics, networks, substance use, risk/prevention behavior, HIV/STI prevalence Chicago
National HIV Behavioral Surveillance (NHBS)	June-December 2017 (MSM cycle V)	Time location venue sampling of White, Black, Hispanic/Latino MSM (transwomen not eligible) of all ages (n ~ 500)	Sociodemographics, substance use, risk/prevention behavior, HIV prevalence, PrEP & ART use, partner by partner characteristics/sex behaviors for up to 3 partners
CDPH HIV surveillance	Ongoing	HIV surveillance records	HIV incidence and prevalence, retention in care, viral suppression
US Census Bureau	Ongoing	Demographics of Chicago population	Age-specific mortality rates, population growth rates, population size overall and by subgroup

Agent-based models for evaluating adaptive interventions

- Adaptive interventions modify intervention or intervention components based on participants' initial response
- ABM can provide insights about the potential impact of adaptive interventions
  - Observe predicted patient trajectories to inform and adjust dose or other intervention components
  - Adjustments to service-level factors, such as provider training
- Can model dynamic, time-varying processes, multicomponent interventions

Agent-based models for implementation science research

- What-if scenarios can be used to evaluate questions at a conceptual level at the beginning of the implementation process even if empirical data are limited
  - Rerun models under different implementation strategies or policy scenarios
  - Can identify barriers and strengths early in the implementation process
- Incorporate behavioral rules at the individual level and organizational-level interactions
- Individual interventions; combinations of interventions applied simultaneously or sequentially; multi-level interventions; cost-benefit analysis