



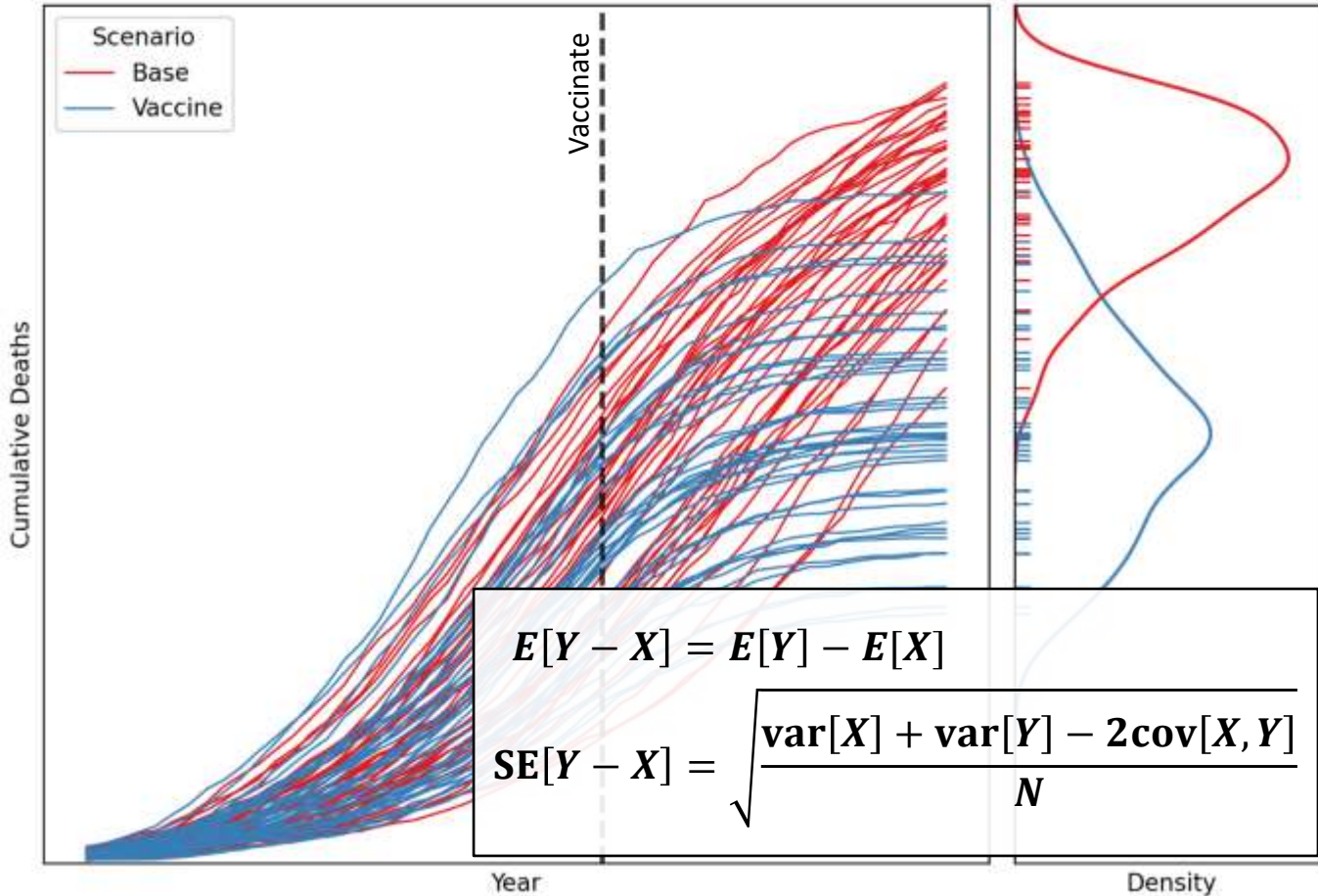
IDM INSTITUTE FOR
DISEASE MODELING

Noise-free comparison of stochastic agent-based simulations using common random numbers

Dan Klein, Romesh Abey Suriya, Robyn Stuart, and Cliff Kerr

BILL & MELINDA
GATES *foundation*

How many deaths are averted by a vaccine?



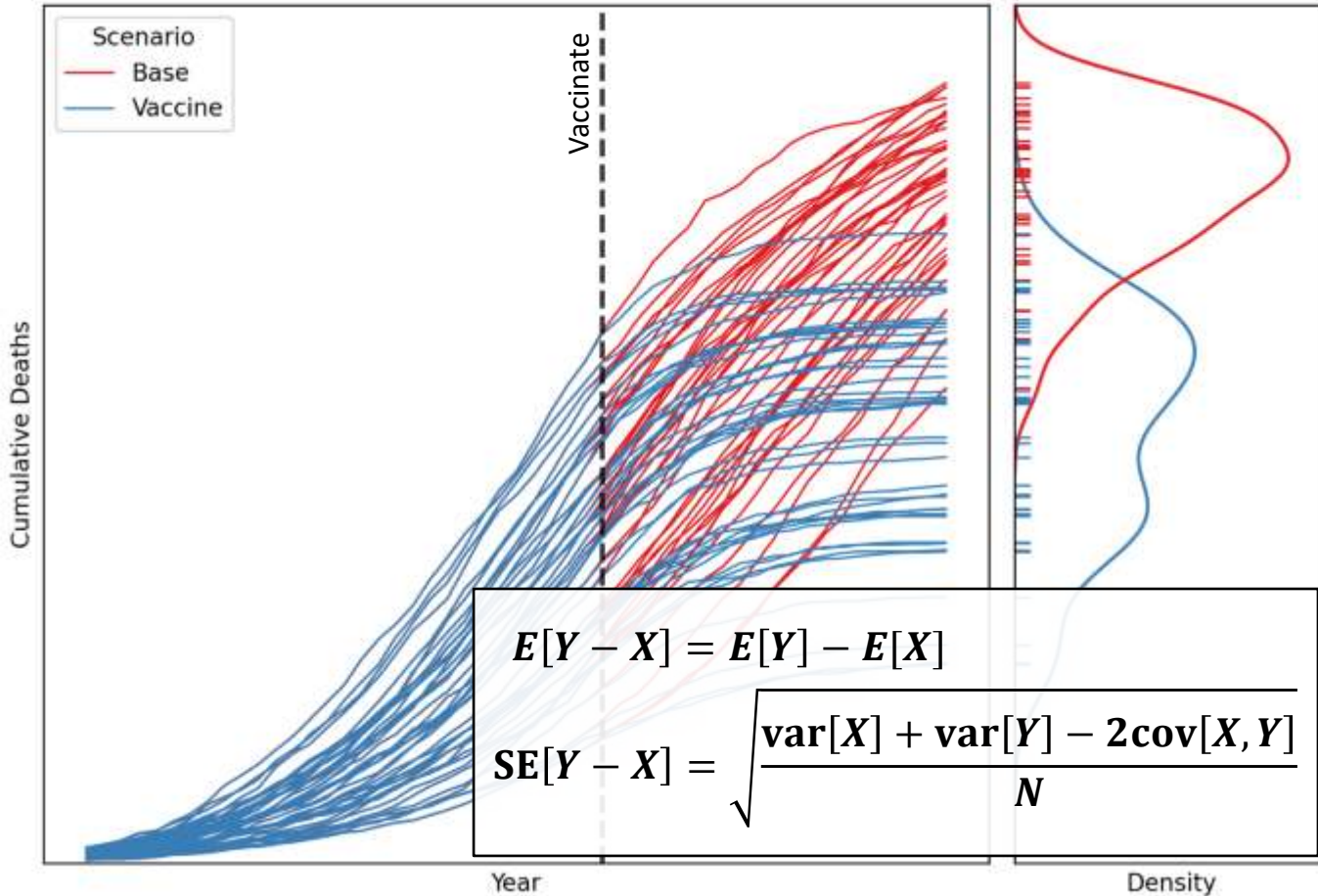
What causes differences between base and vaccine simulations?

1. Mechanistic effects
2. Different populations
3. Unwanted noise due to random numbers

Can we do better?

- Increase covariance (common seeds)
- Common random numbers eliminate unwanted noise

How many deaths are averted by a vaccine?



What causes differences between base and vaccine simulations?

1. Mechanistic effects
2. Different populations
3. Unwanted noise due to random numbers

Can we do better?

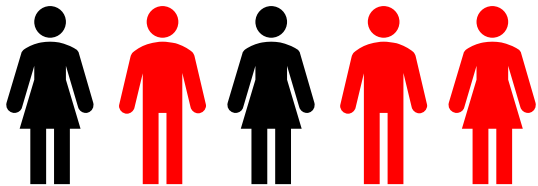
- Increase covariance (common seeds)
- Common random numbers eliminate unwanted noise

Fundamentally, ABMs don't work like you'd assume

Pseudo-random numbers, e.g. `np.random.randint(10)`

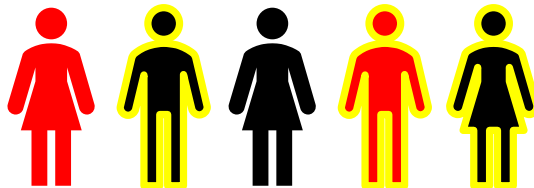


Infect if $r < 5$



Vaccinate if $r < 5$

Infect if $r < 4$ if v_x , 5 otherwise



Problem: Centralized random number stream used for everything. Any difference between simulations results in *misaligned random draws*.

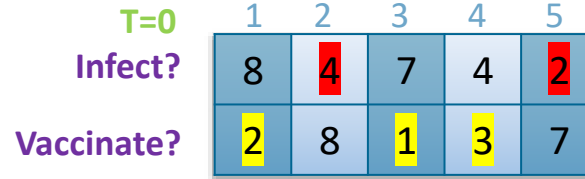
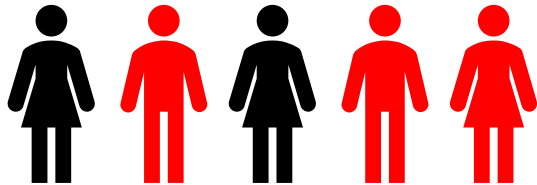
Result: Outputs comparable only at the population level, not the individual level – even with common seeds.

Solution is Common Random Numbers (CRN)

- Idea dates back to the early days of Monte Carlo simulation
- Use the same random draw per **decision**, **agent**, **time**
- Simple idea, but **CRN has never been achieved in ABM** (until now)

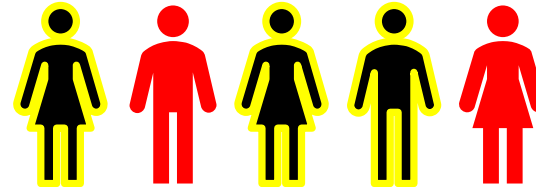


Infect if $r < 5$



Vaccinate if $r < 5$

Infect if: $r < 4$ if v_x , $r < 5$ otherwise



CHALLENGES:

1. Births
2. Interactions

CRN: Differences now due to action of vaccine at individual level, not random number noise

Modeling innovations required to achieve CRN

1. Separate random streams for each **decision**, seed offset from hash of label
2. On **time** step t_i , reset each stream and “jump” t_i times
3. Each new **agent** is assigned a “slot” by a parent, used as index in draws array

Example

Agent 0 has slot 0, gets 8

Agent 1 has slot 4, gets 2

Index	0	1	2	3	4
Draws	8	4	7	4	2

4. Pairwise random numbers by bit-mixing per-agent random numbers

$$r_{ij} = \text{XOR}(r_i * r_j, r_i - r_j) / \text{max}(\text{uint64})$$

Use pairwise draws for interactions like networks and transmission

We have implemented CRN in Starsim

- CRN methods implemented in the new Starsim ABM Framework
- User-centric design
- Straightforward statistical distributions internally implement pseudo-random decision streams, jumping, and slot-based slicing

Does it work? Does it matter?

```
import starsim as ss

class SIR(ss.Infection):
    def __init__(self, user_pars=None):
        self.default_pars(
            dur_inf = ss.weibull(c=3, scale=10)
        )
        self.update_pars(user_pars)
        return

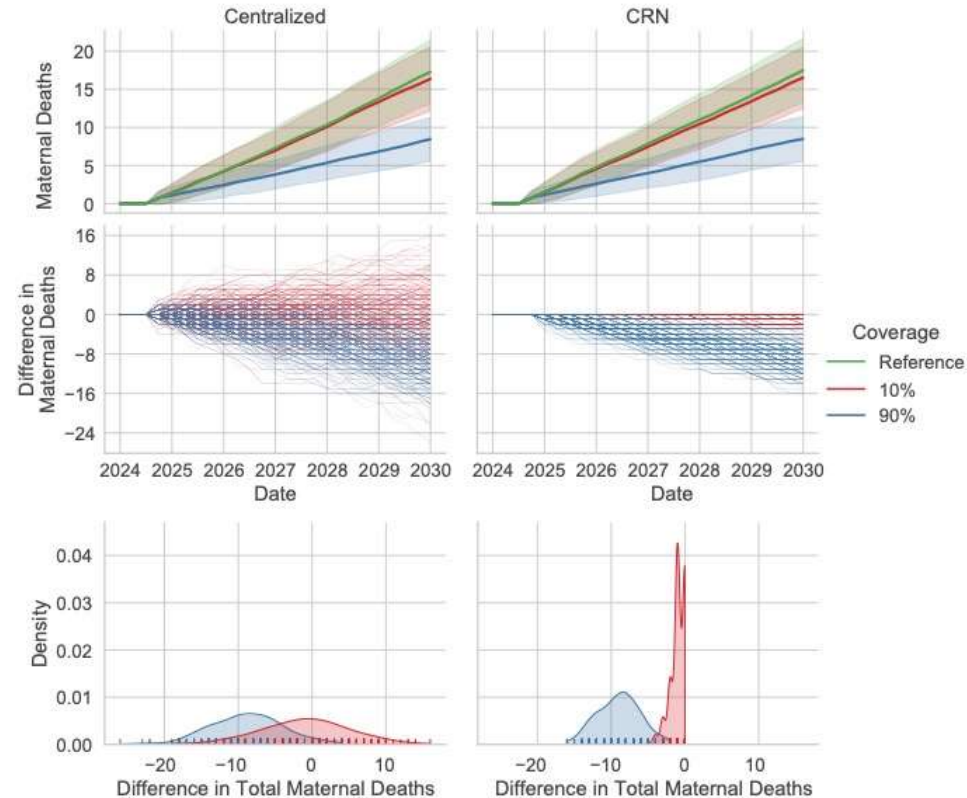
    def set_prognoses(self, uids):
        duration = self.dur_inf.rvs(uids)
        self.t_recovery = self.sim.t + duration
        return
```

MNCH: Prevention of Postpartum Hemorrhage

- Births and deaths, no interactions
- MMR in sub-Saharan Africa estimated at 500 per 100,000
 - PPH is a leading cause, about 25%
 - E-MOTIVE package efficacy ~60%
 - Increased infant mortality for orphans
- N=100,000 agents, 250 replicates
- Differences by pairing random seeds

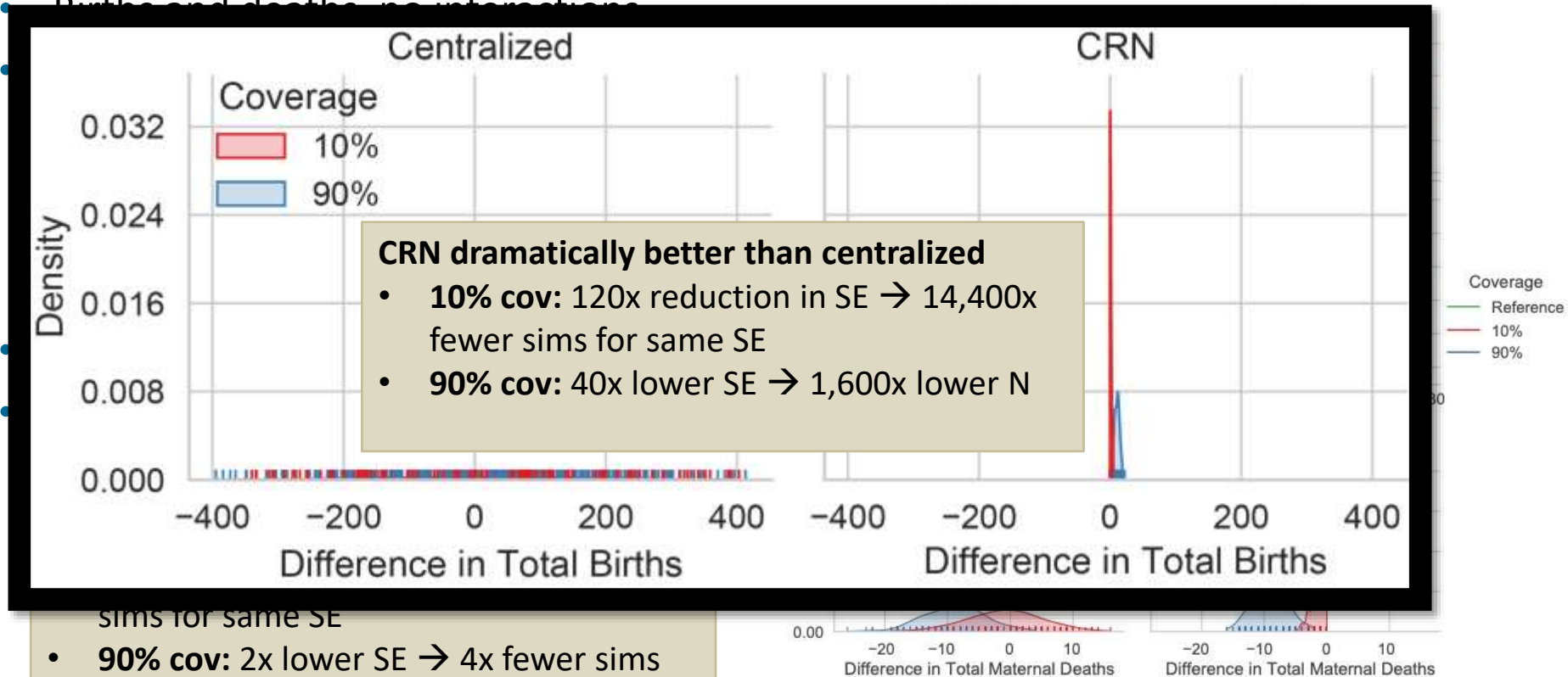
CRN dramatically better than centralized

- E-MOTIVE never worse with CRN
- **10% cov:** 6x reduction in SE → 36x fewer sims for same SE
- **90% cov:** 2x lower SE → 4x fewer sims



MNCH: Prevention of Postpartum Hemorrhage

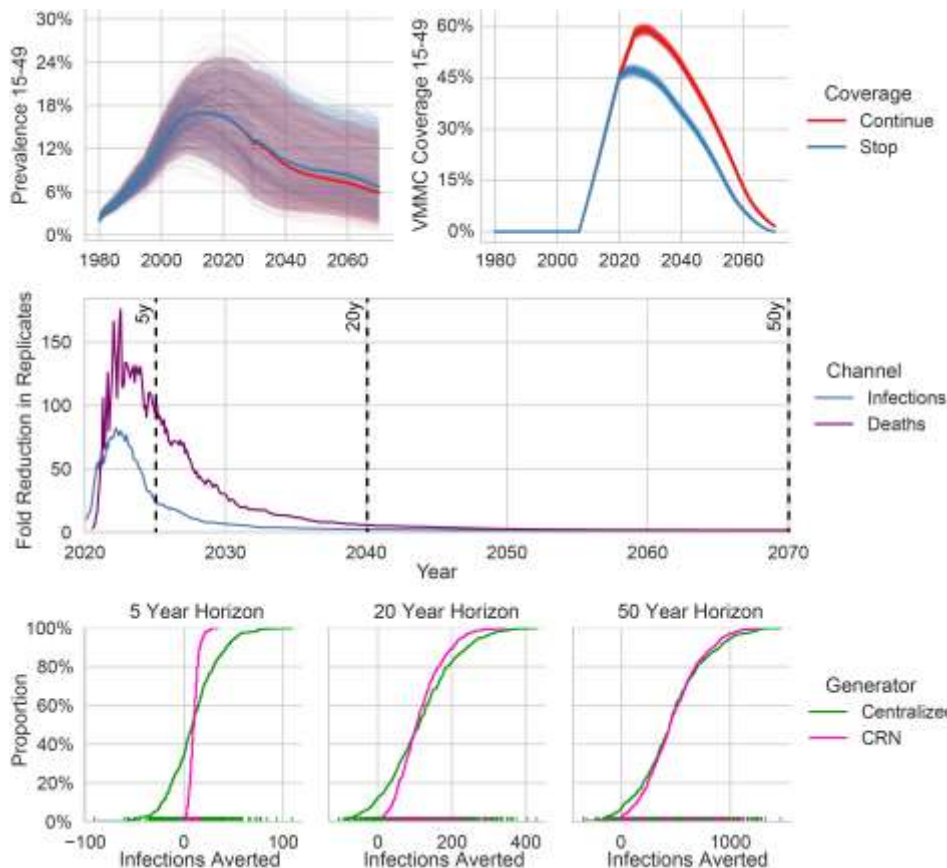
Birth and death rate interactions



sims for same SE

- **90% cov:** 2x lower SE → 4x fewer sims

HIV: Impact of continued VMMC scale-up



Full dynamic transmission model

- Inspired by [HIV modeling consortium](#)
- Evaluation periods of 5y, 20y, and 50y
- Basic HIV module including prognosis, ART, and transmission
- SSA-like, 10k agents, 250 reps
- VMMC efficacy 60%, continue or stop

Sim savings	5y	20y	50y
Infections Averted	24x	2.3x	1.4x
Deaths Averted	95x	5.7x	1.6x

Significant savings at 5y, but benefits decrease with longer evaluation periods

CRN is a significant achievement for ABM

- Method overcomes a longstanding signal-to-noise problem in ABM
- A first for comp epi, enabled by multiple innovations
- User-friendly implementation
- **Benefits include:**
 - Fewer simulations for same SE
 - Results interpretable at individual level
 - Aids scientific communication
 - Scenarios, sensitivity, & calibration
- **Find:** Value depends on use case
 - *New standard* for non-interacting & high-sensitivity applications
 - Less benefit for large perturbations
- **Limitation:** population scaling
 - Random numbers are cheap
 - But pairwise algos are $O(N^2)$
- Currently socializing methods & applying where appropriate
- **Publication:** D. Klein, R. Abeyesuriya, R. Stuart, and C. Kerr, “[Noise-free comparison of stochastic agent-based simulations using common random numbers](#)” Submitted to PLOS CB