MIND THE AGE GAP

A simulation study of age-mixing patterns and HIV incidence (preliminary results)

Wim Delva

Jori Liesenborgs

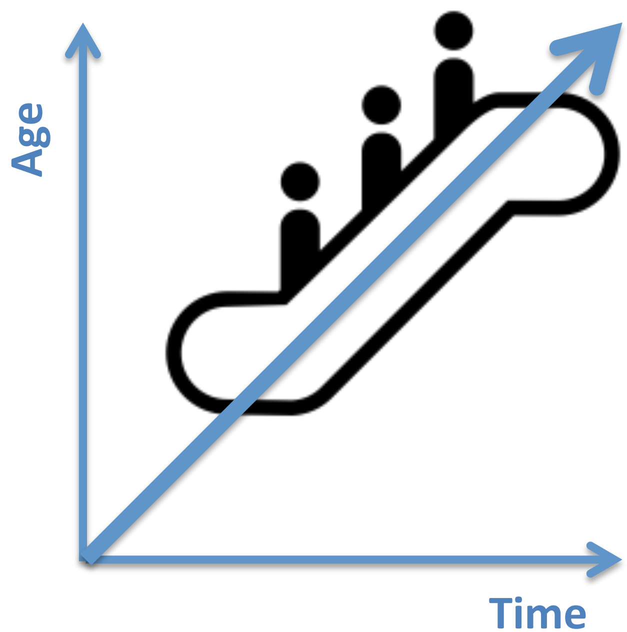
Stéphane Helleringer

Niel Hens

SIMPACT

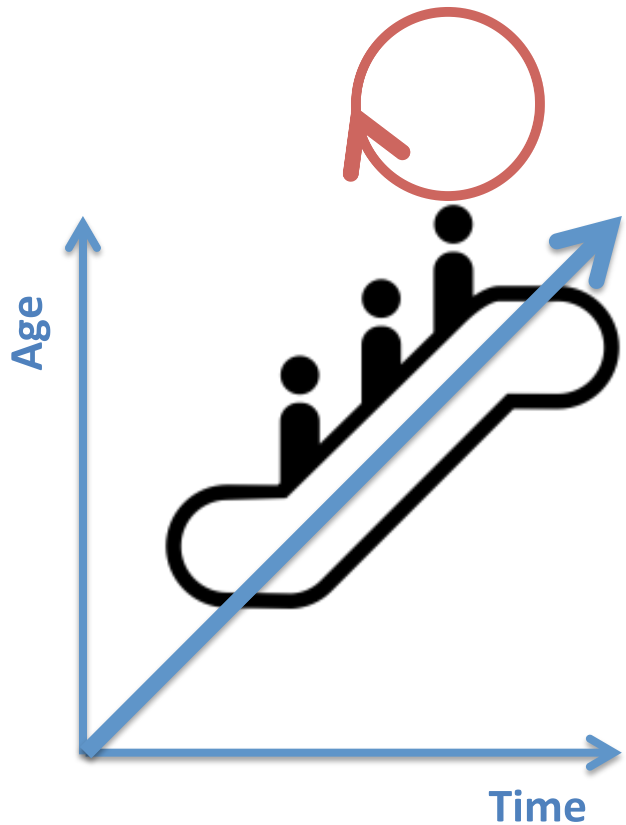
Why might ageing matter?

- Age and time are (perfectly) correlated



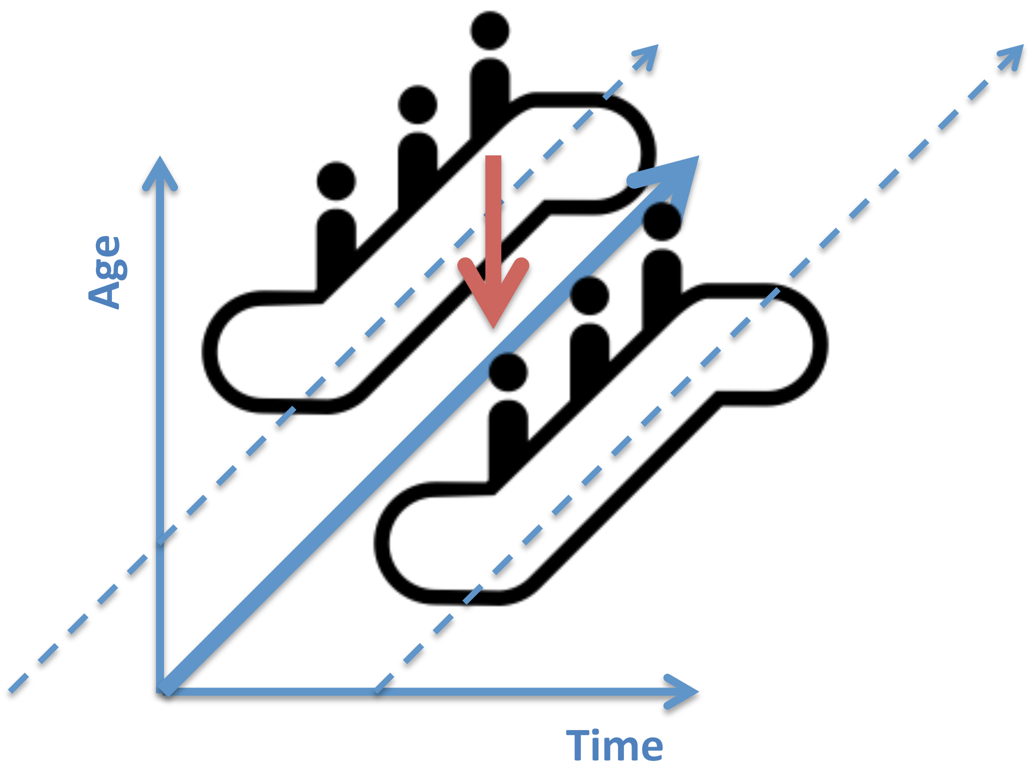
Why might ageing matter?

- Perfect age-assortativity is a trap for HIV



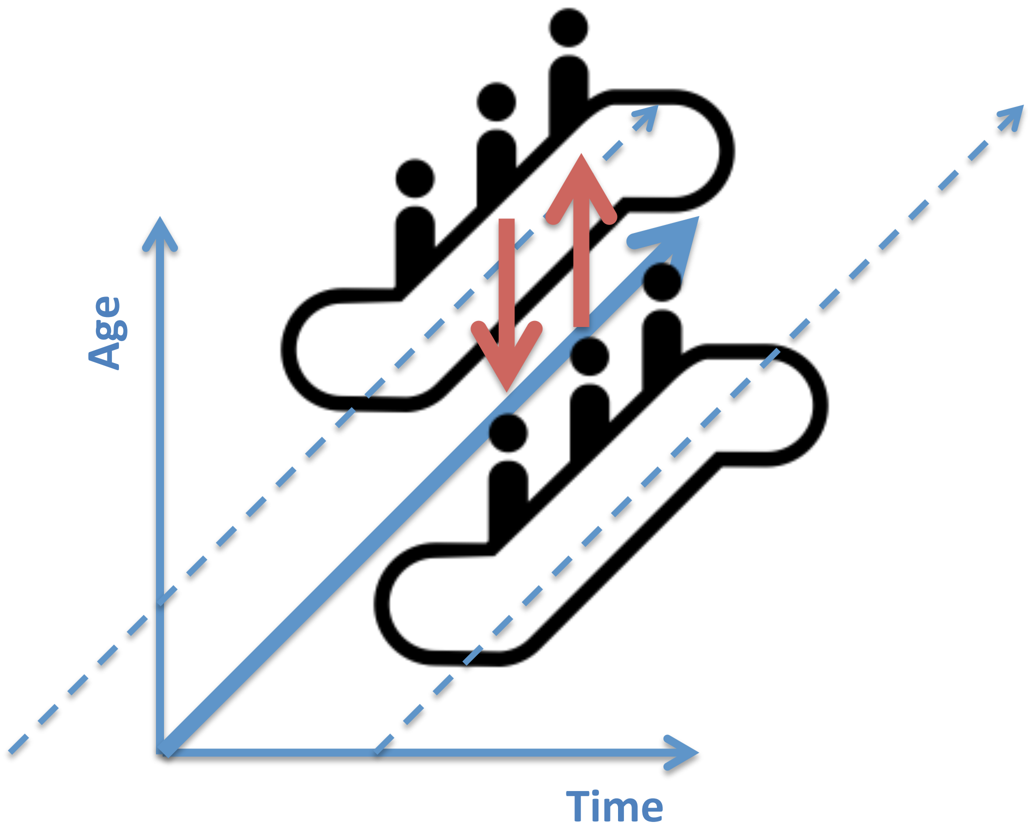
Why might ageing matter?

- HIV needs to find ways to stay "rejuvenated"



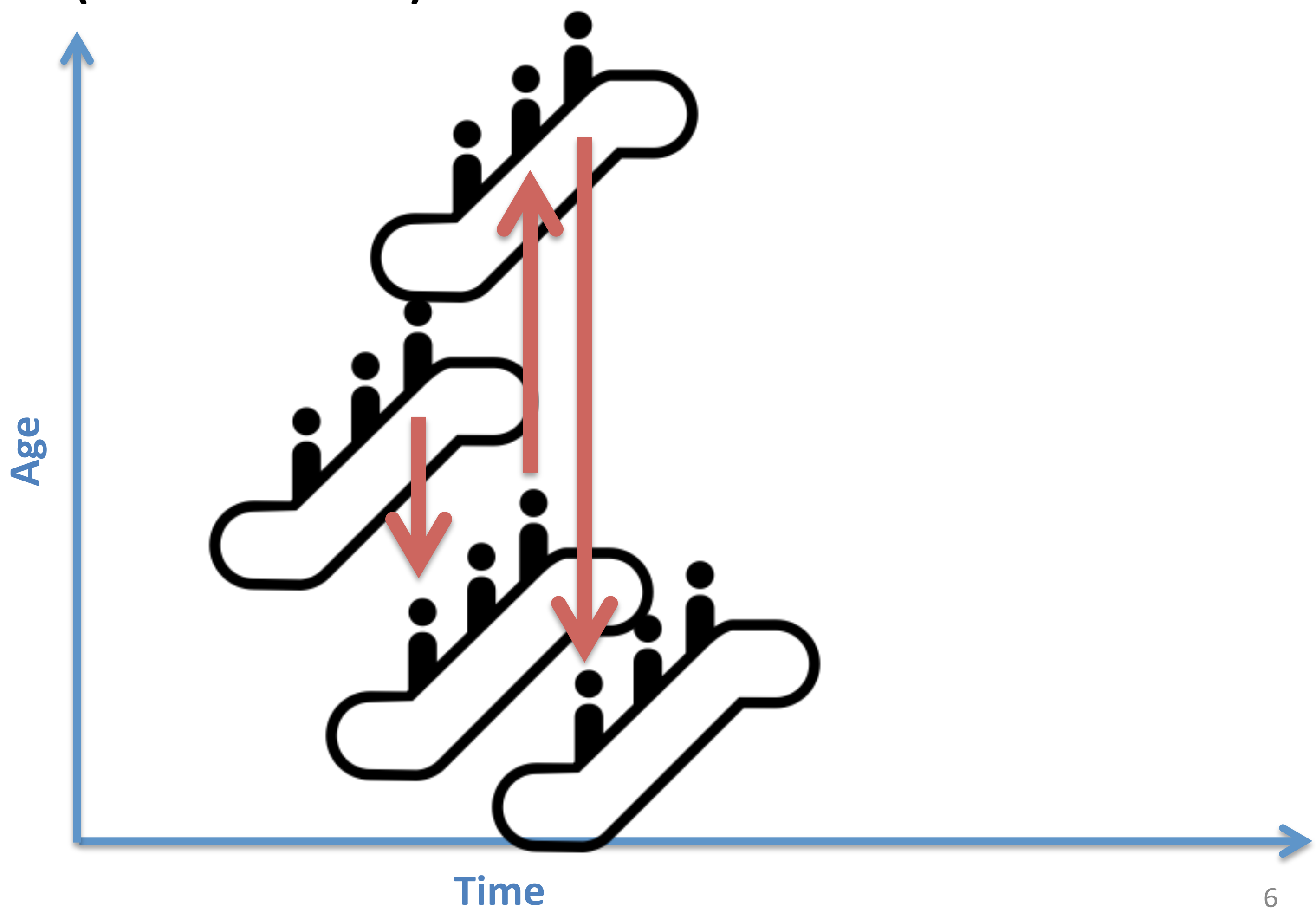
Why might ageing matter?

- A large, **fixed** age gap is not sufficient



Why might ageing matter?

- HIV needs variation of age gaps within a person's (infectious) lifetime



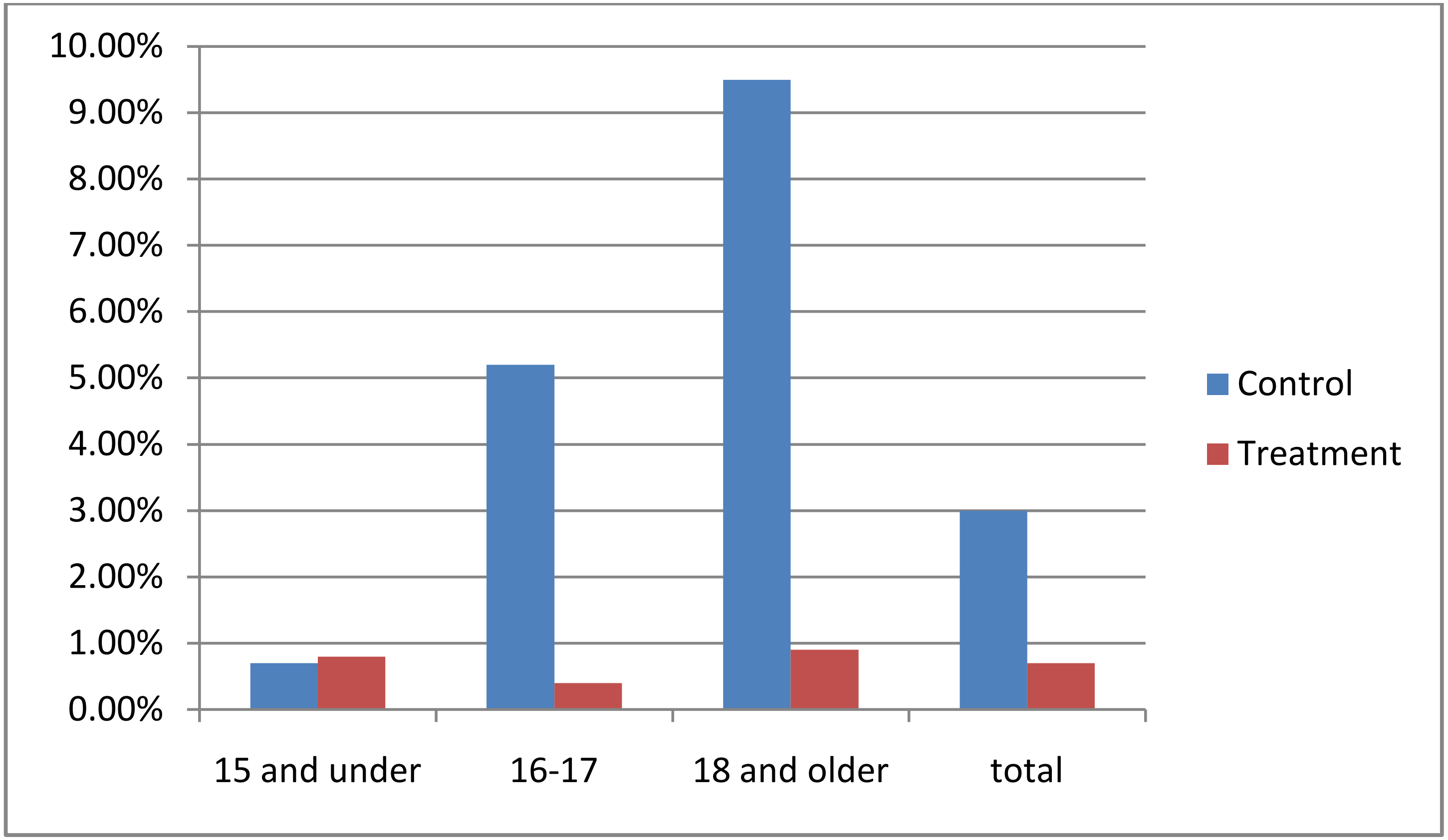
The Malawi cash transfer trial

- Community-randomised controlled trial
- Cash transfer to girls, their families and schools
- No direct targeting of sexual behaviours
- Intervention group: More likely to stay in school, less likely to have partner of 25yo or older, less likely to have transactional sex
- Weighted HIV prevalence at 18mo: 1.2% vs. 3.0%

Baird et al. Lancet 2012

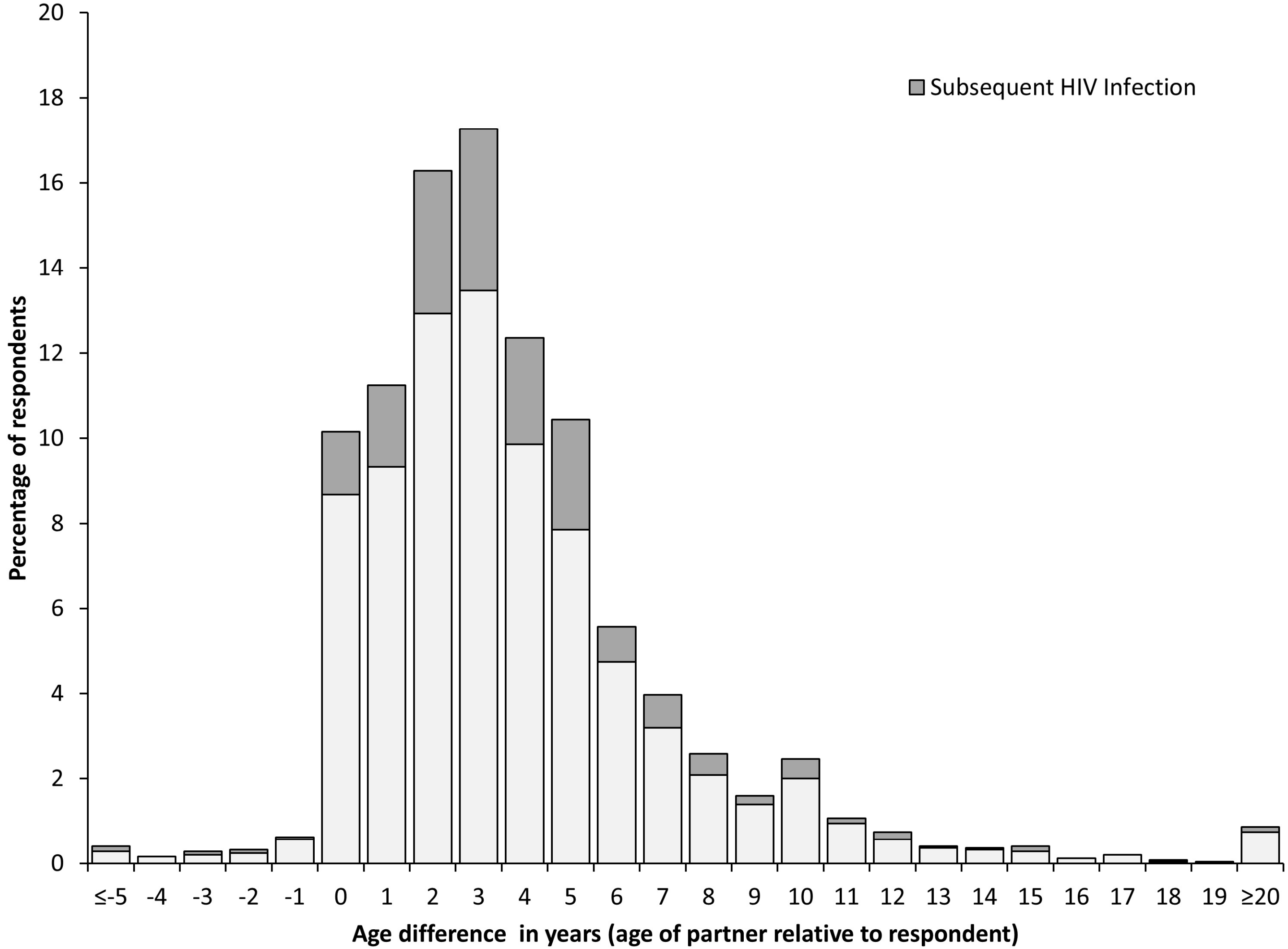
The Malawi cash transfer trial

HSV-2 prevalence at follow-up



Baird et al. 2012

Individual-level analysis

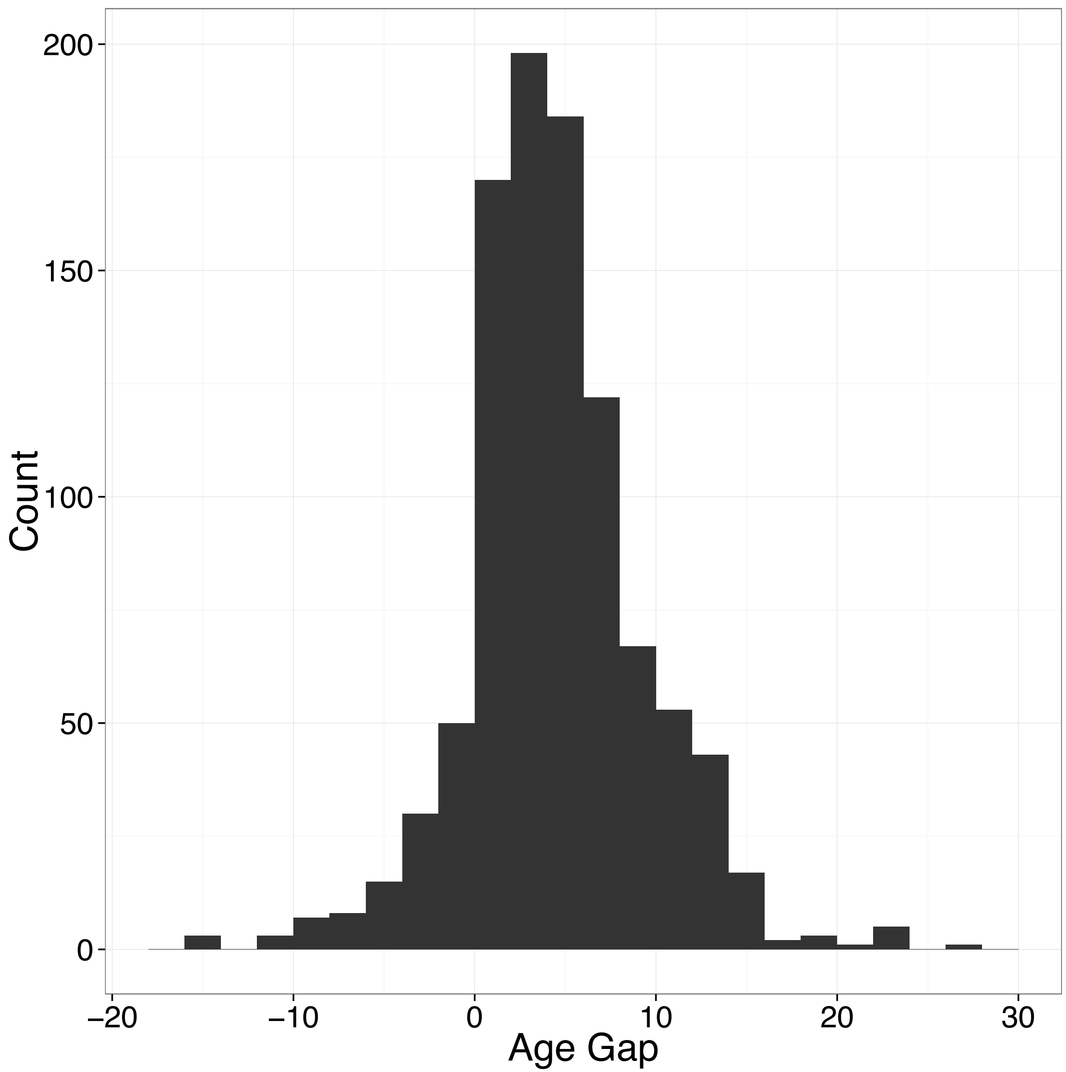


Harling et al. JAIDS 2014

Issues to be addressed

- Effect of (*altering)* age-mixing pattern on HIV incidence, over time [and by gender and age]
- Importance of hypersusceptibility among young women?
- Individual-level (short-term) versus population-level (long-term) effects of age-mixing pattern

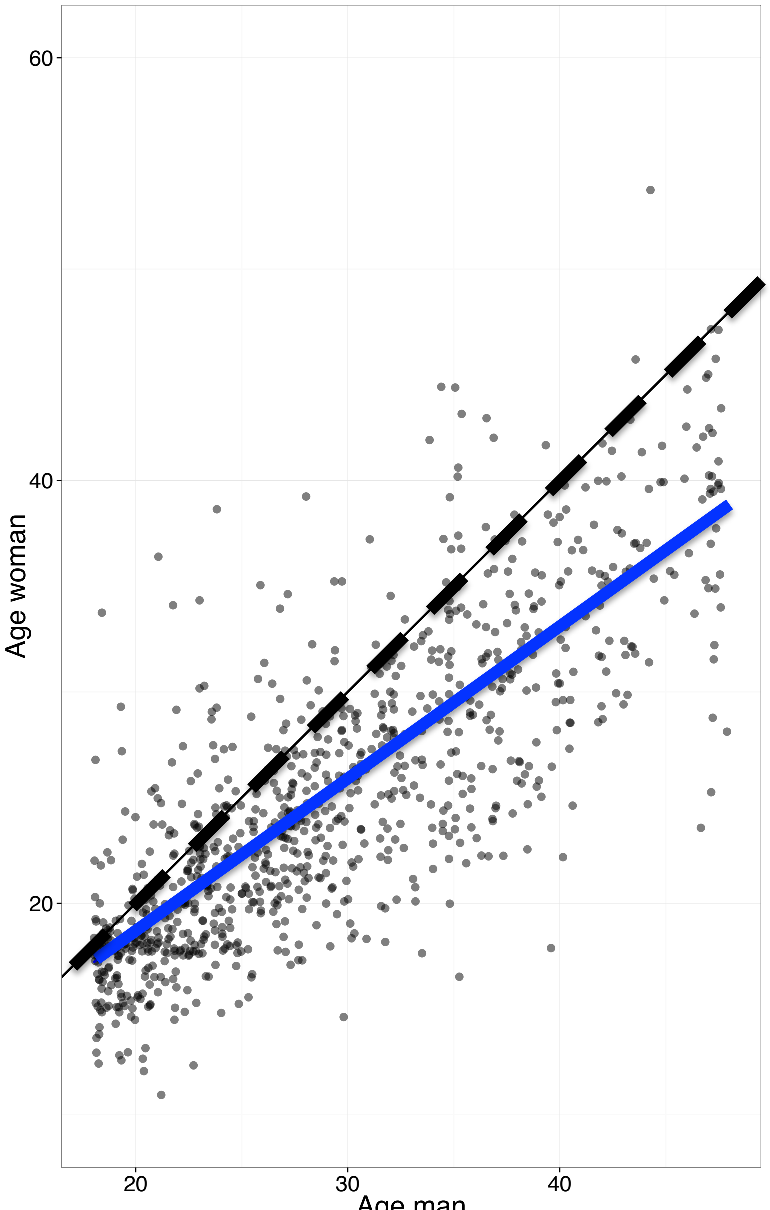
Age-mixing pattern on Likoma Island, Malawi



Average age gap: 4.0 years

Overall SD of age gaps: 4.9 years

Age-mixing pattern on Likoma Island, Malawi



Slope:
0.72 [0.69 – 0.75]

Between-subject SD:
1.5 years [1.1 – 2.0]

Within-subject SD *at 18 yo:*
2.3 years [2.0 – 2.6]

Within-subject SD *at 49 yo:*
5.3 years

Power (W-s **Variance**):
0.24 [0.19 – 0.30]

Why Agent-based models?

- Autonomy: agents make “their own decisions” based on agent-specific and system state variables
- Heterogeneity: agents don’t all behave in the same way
- The system is not memoryless (e.g. HIV-mortality rate increase with time since infection)

Events in Simpact

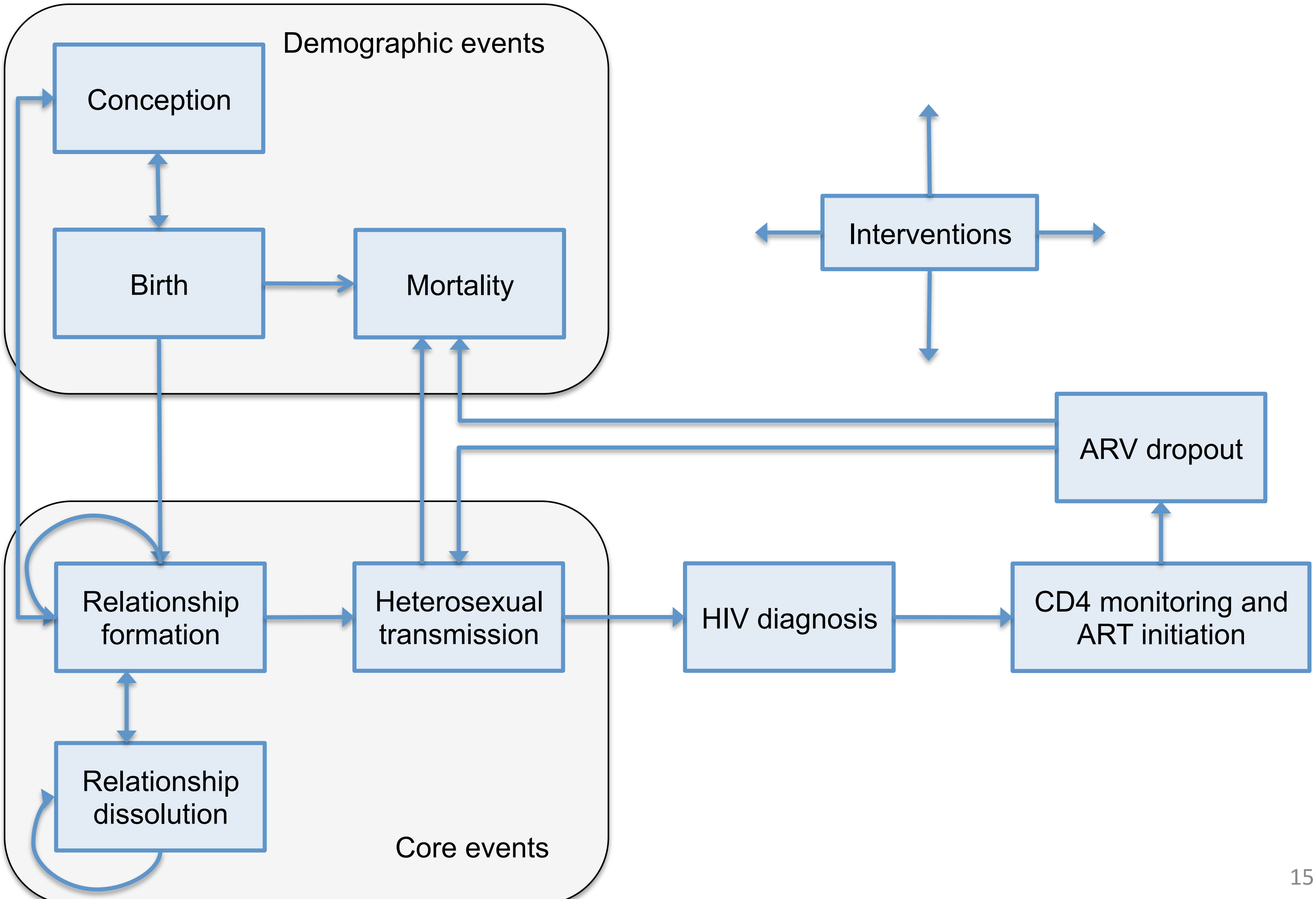
1. Relationships formation
2. Relationship dissolution
3. HIV transmission
4. Mortality
5. AIDS mortality
6. Conception
7. HIV Diagnosis
8. ART treatment dropout

9. HIV seeding
10. Intervention

11. AIDS stage
12. Birth
13. Chronic HIV stage
14. Sexual debut
15. HIV infection monitoring

16. Periodic logging
17. Synchronize population statistics
18. Synchronize reference year

Events in Simpact



Hazard function for relationship formation

$h_{F_{ij}}(x,t) = \exp($

$a0 +$

$a1\ (\text{eagerness}_i + \text{eagerness}_j) +$

$a2\ |\text{eagerness}_i - \text{eagerness}_j| +$

$a3\ \text{partners}_i +$

$a4\ \text{partners}_j +$

$a5\ |\text{partners}_i - \text{partners}_j| +$

$\mathbf{a6}\ (t - (tB_i + tB_j)/2) +$

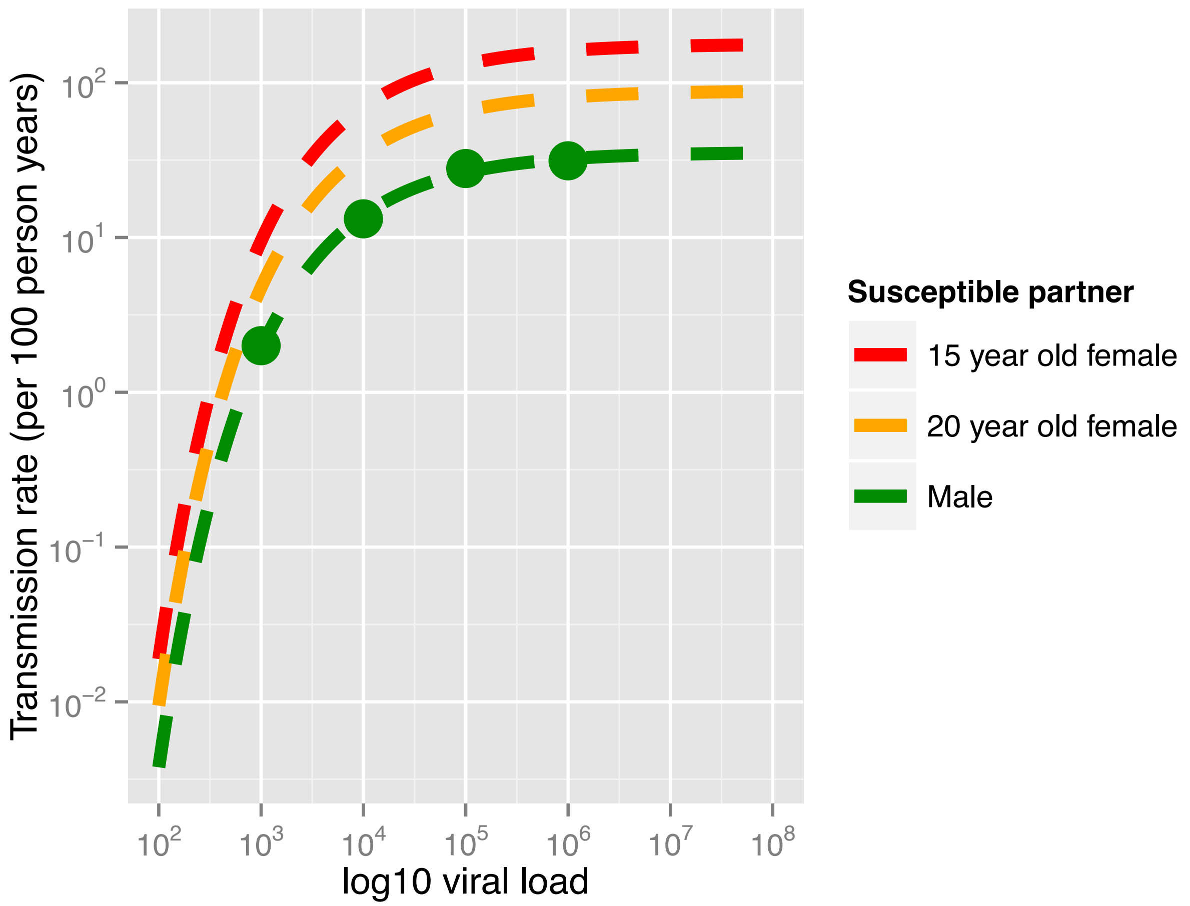
$\mathbf{a7}\ |(a8 - 1)\ tB_i + tB_j - Dp_i - \mathbf{a8}\ t| +$

$\mathbf{a9}\ |(a10 - 1)\ tB_j + tB_i - Dp_j - \mathbf{a10}\ t| +$

$b\ (t - t_r)$

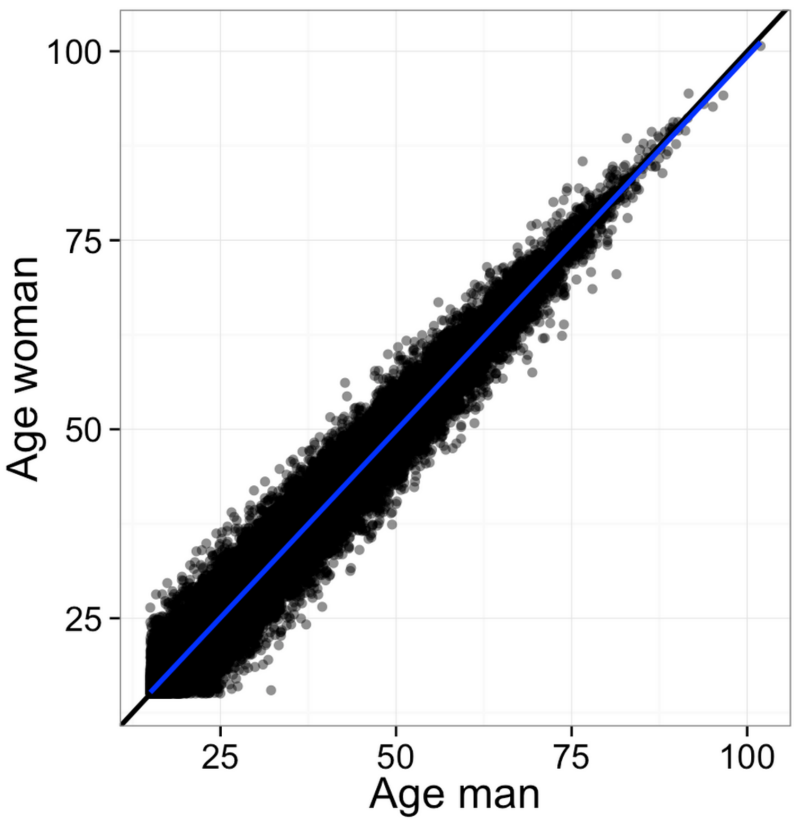
$)$

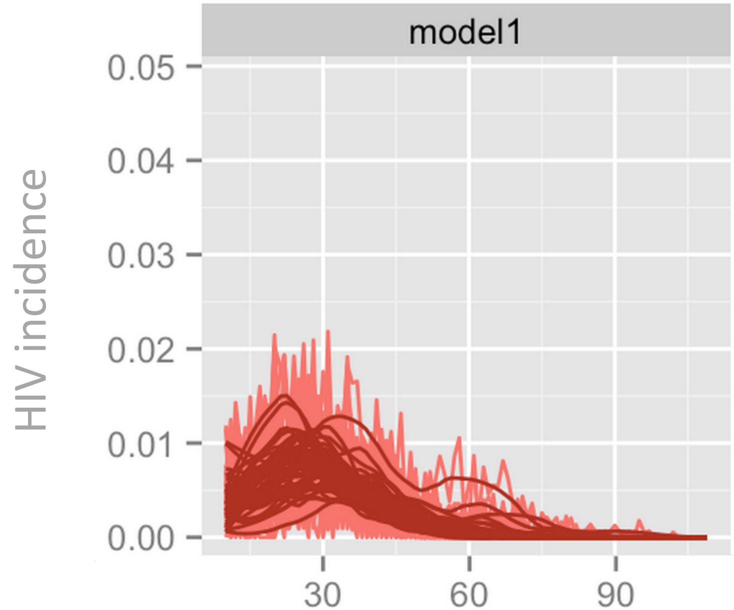
Hazard function for HIV transmission

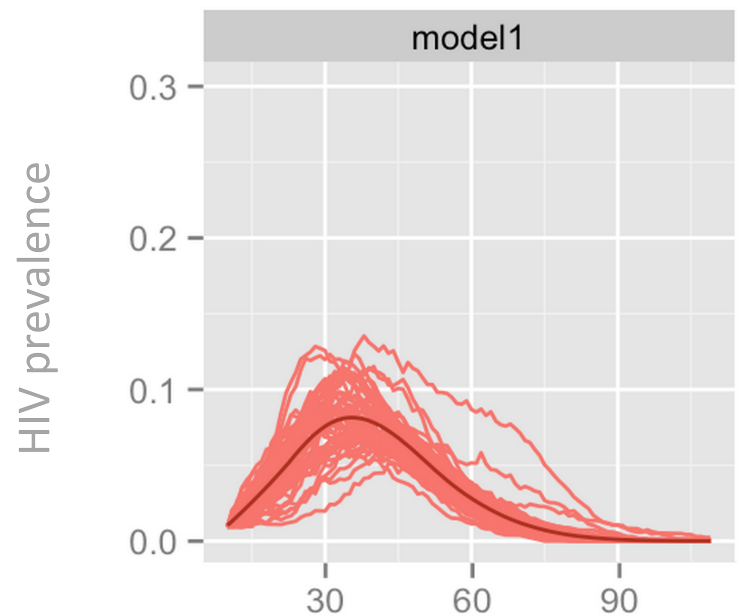


$$h_{T_{ij}}(x,t) = \exp($$
$$a +$$
$$b\, V^{-c} +$$
$$W\, f_1 \exp(f_2\,(A_{w\_ry} - A_{w\_debut}))$$
$$)$$

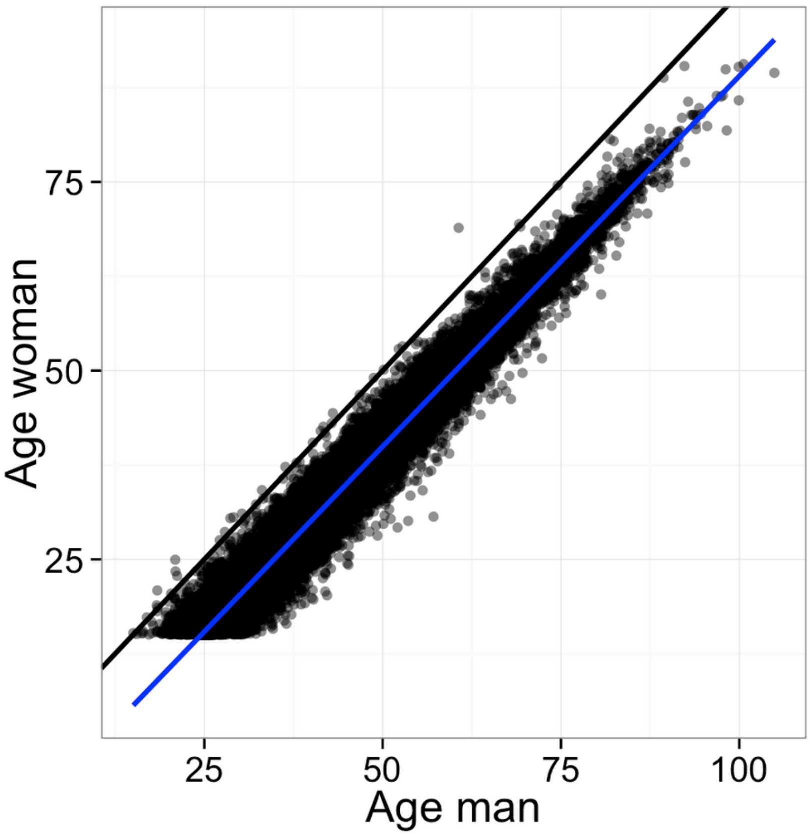
First, some toy examples

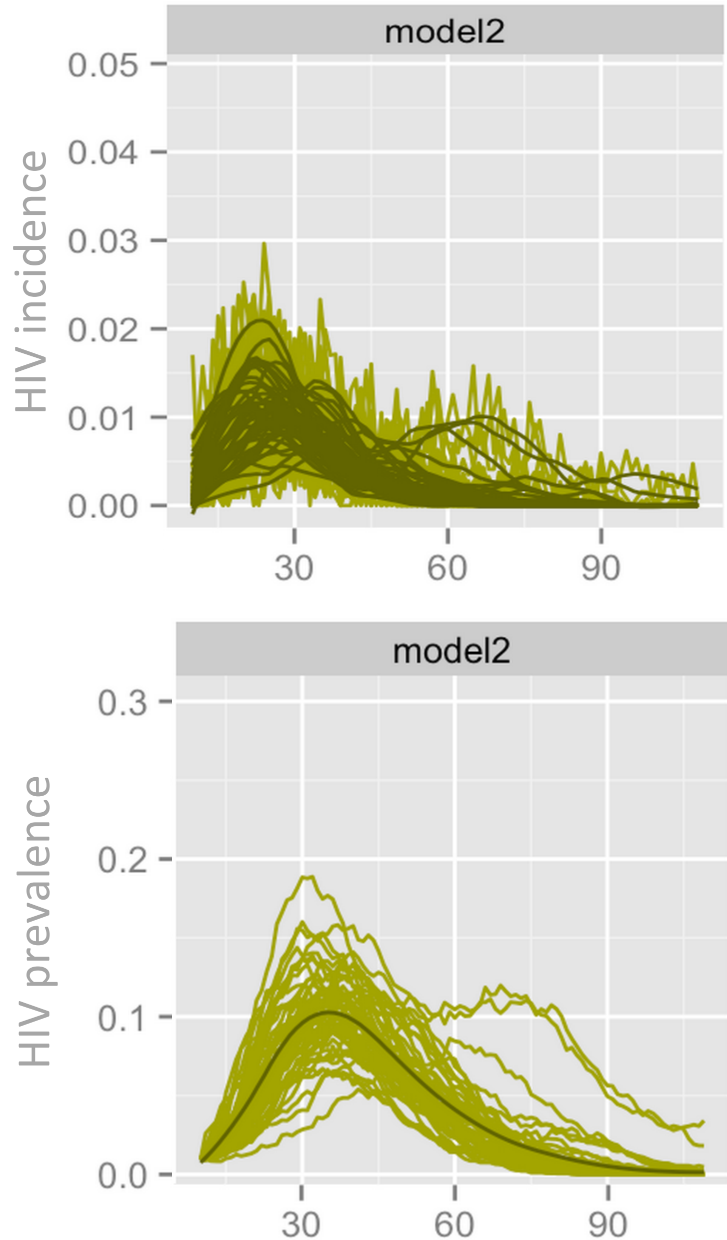




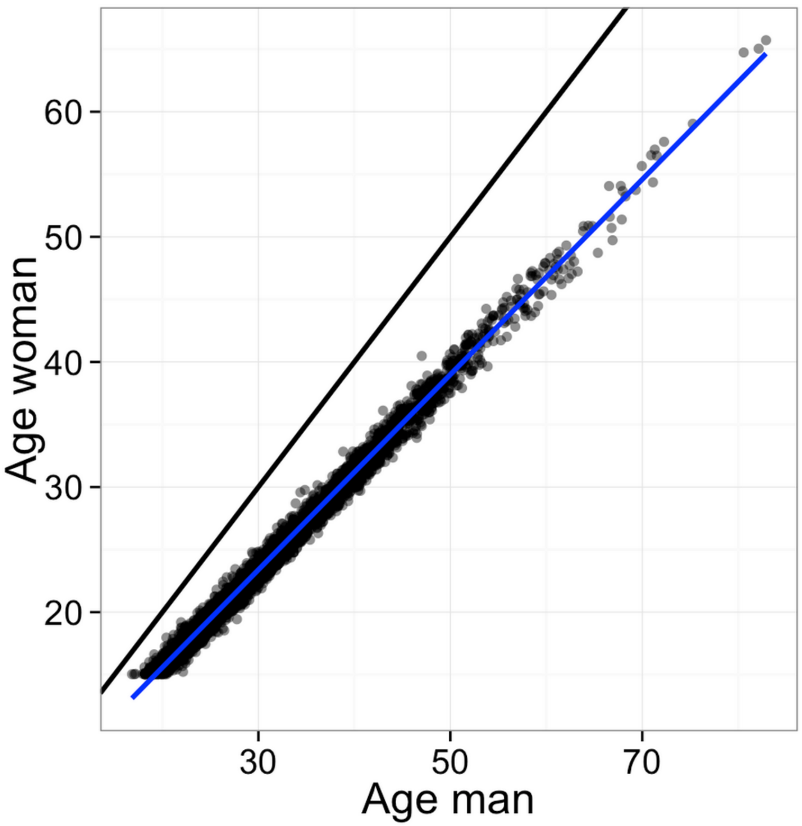


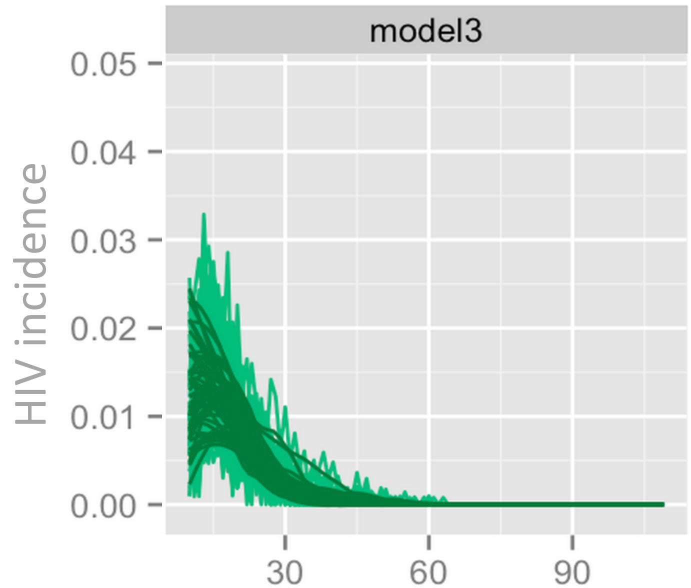
First, some toy examples

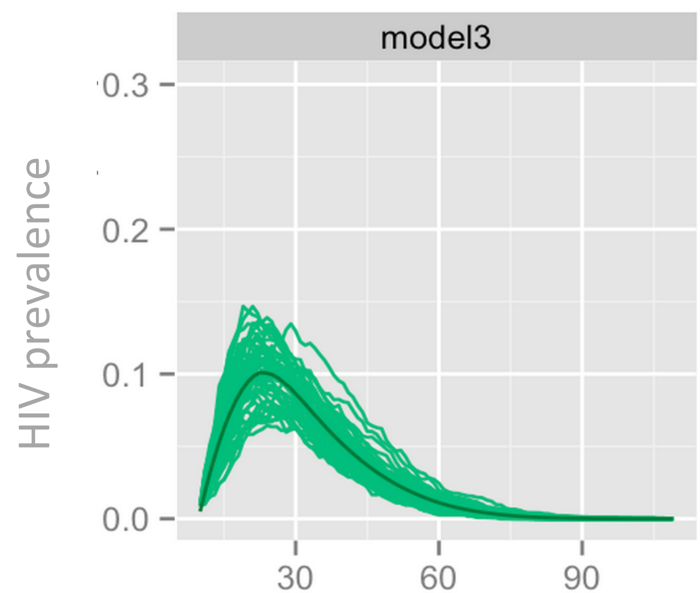




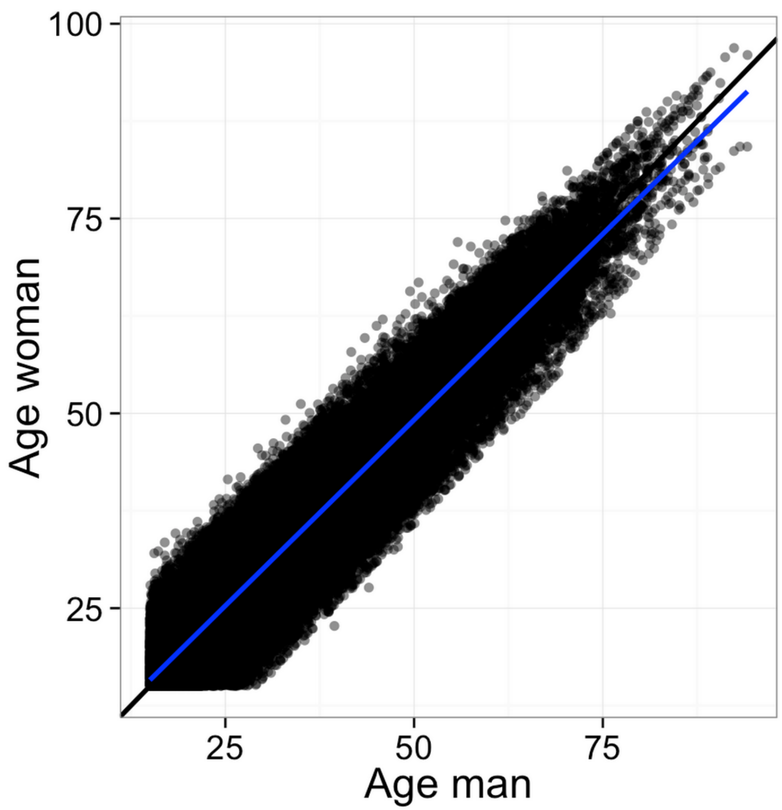
First, some toy examples

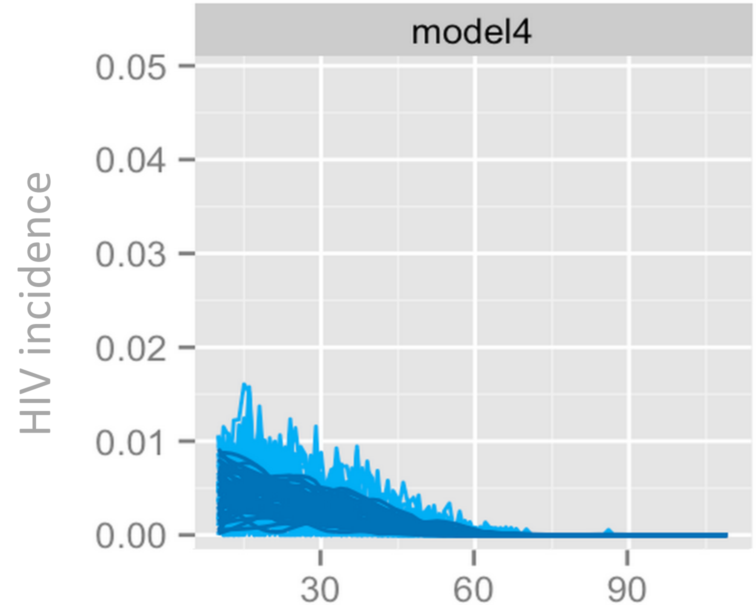


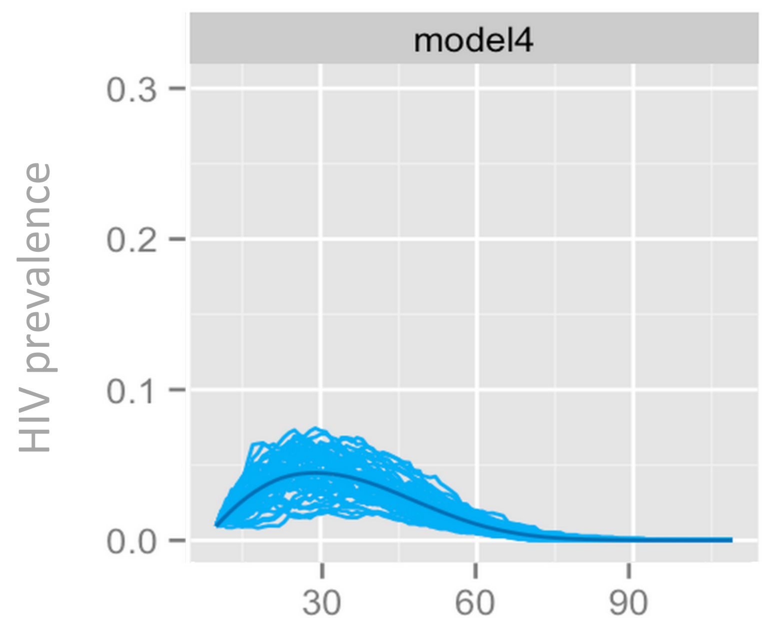




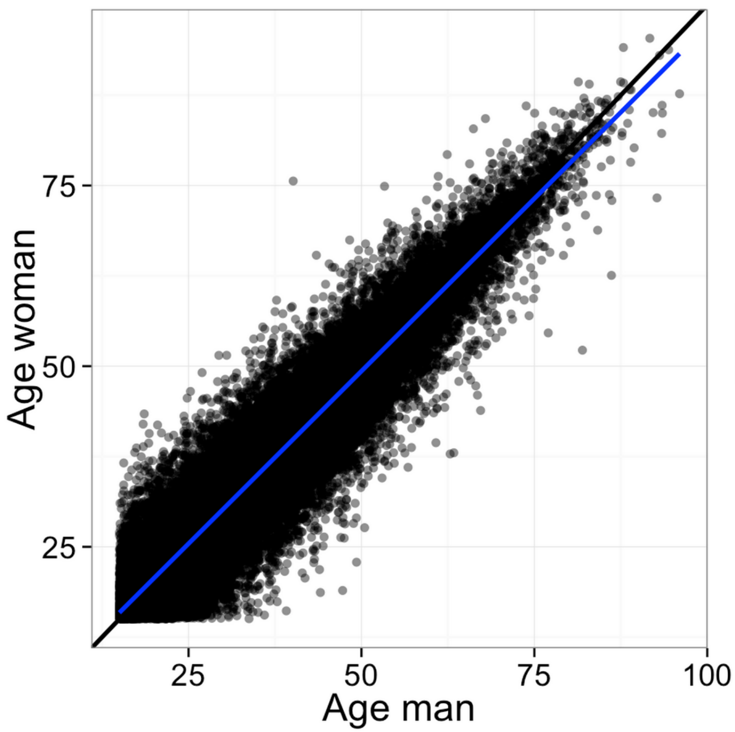
First, some toy examples

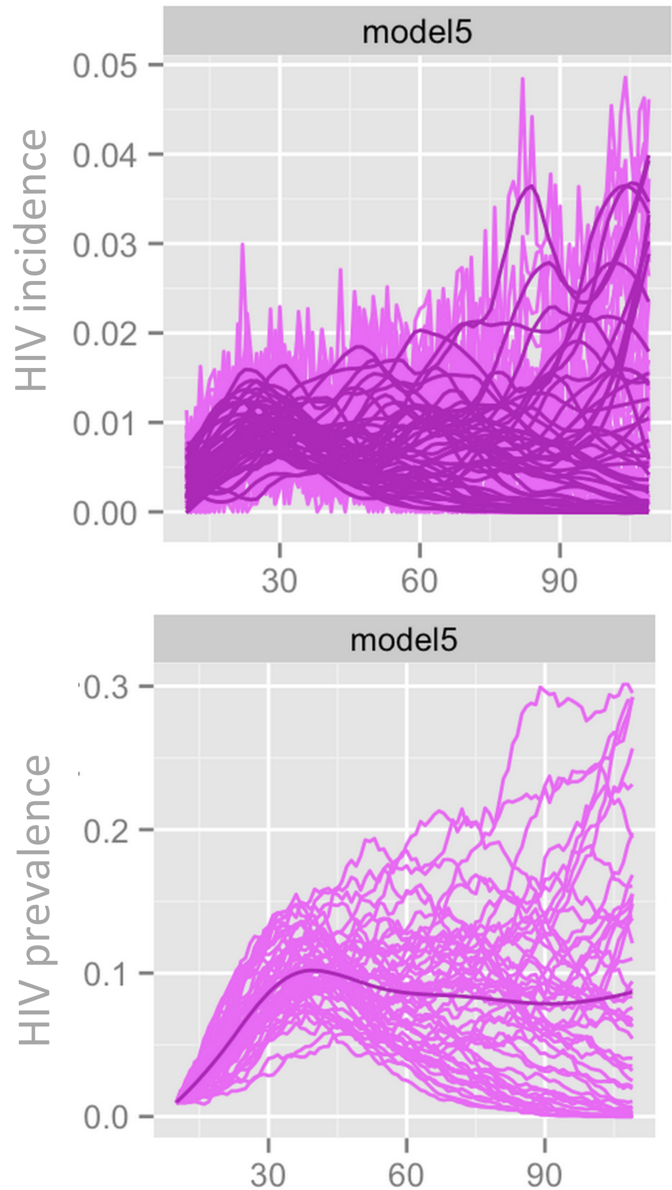






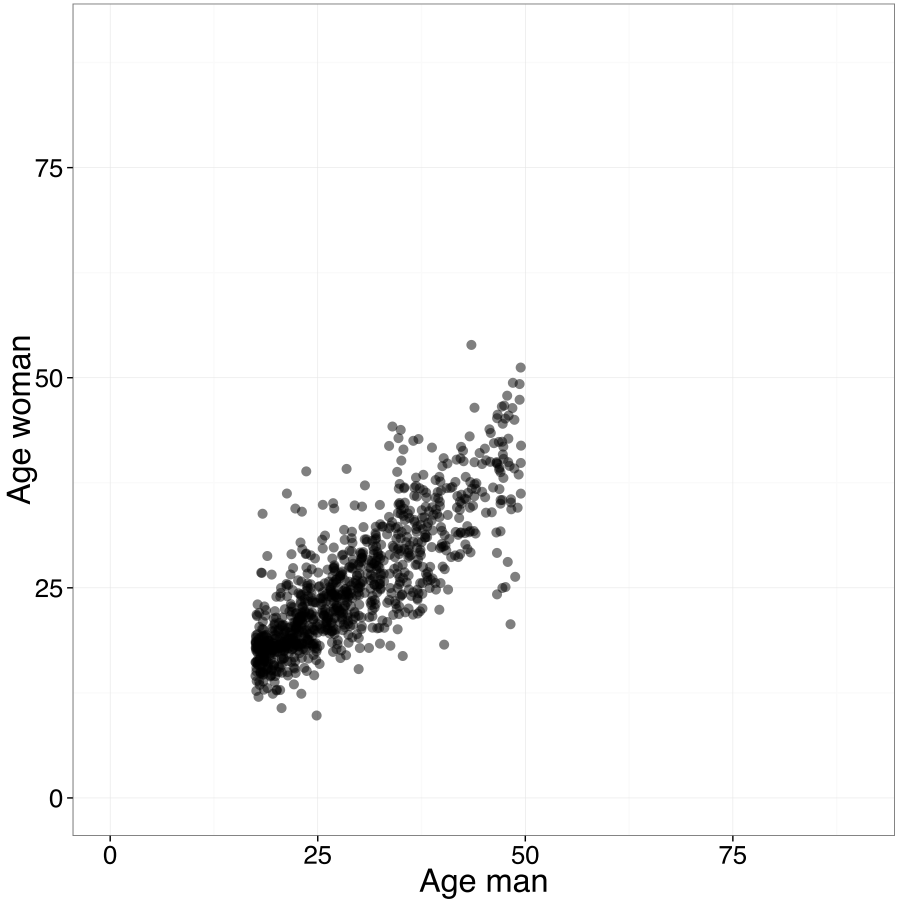
First, some toy examples

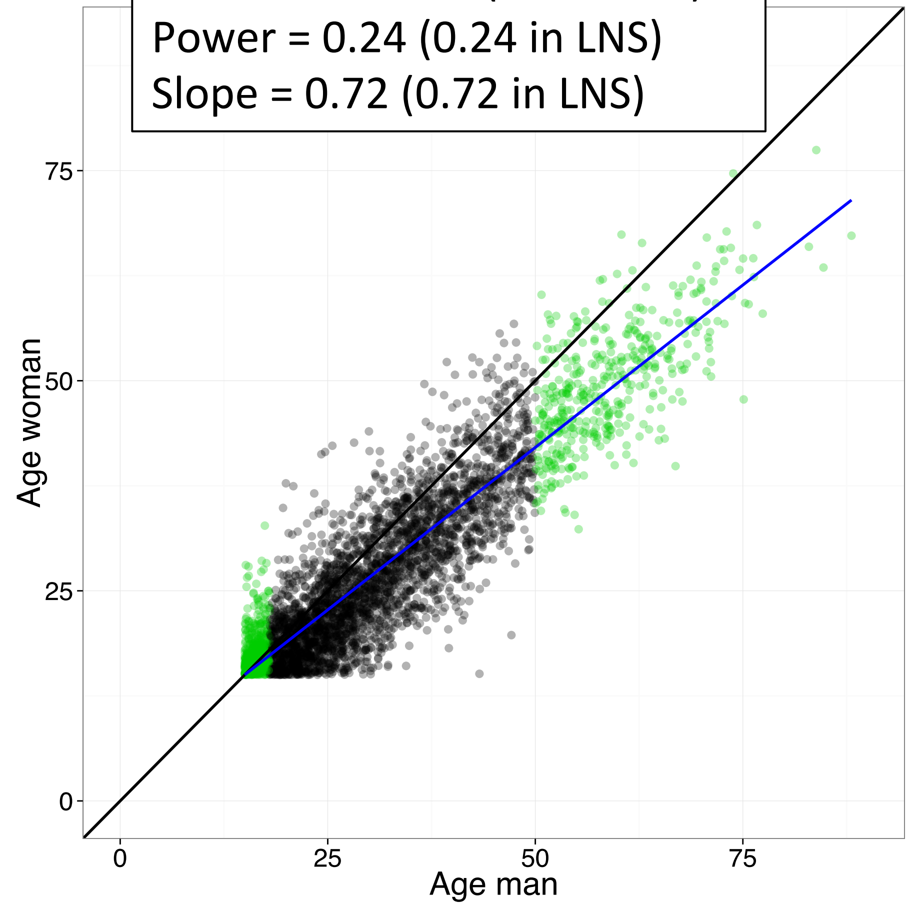




The Reference model (~LNS data)

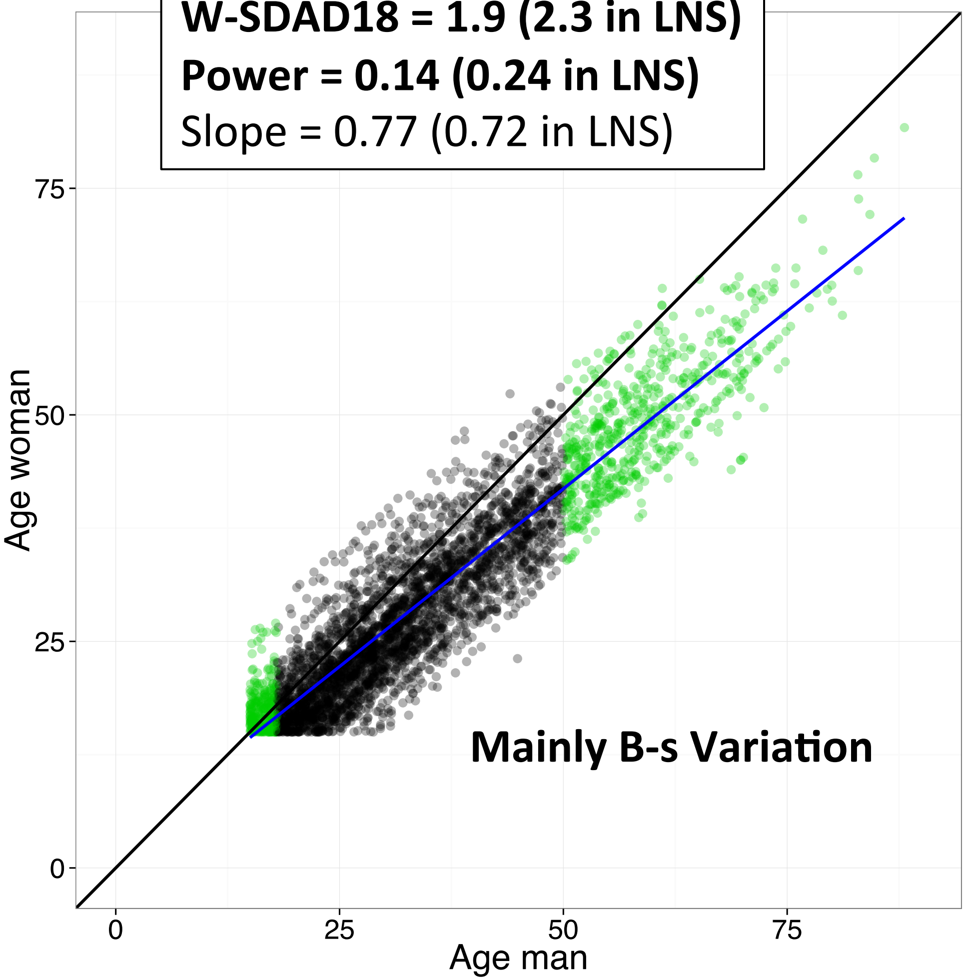
AAD = 3.9 (4.0 in LNS)
SDAD = 5.6 (4.9 in LNS)
B-SDAD = 1.7 (1.5 in LNS)
W-SDAD18 = 2.3 (2.3 in LNS)
Power = 0.24 (0.24 in LNS)
Slope = 0.72 (0.72 in LNS)



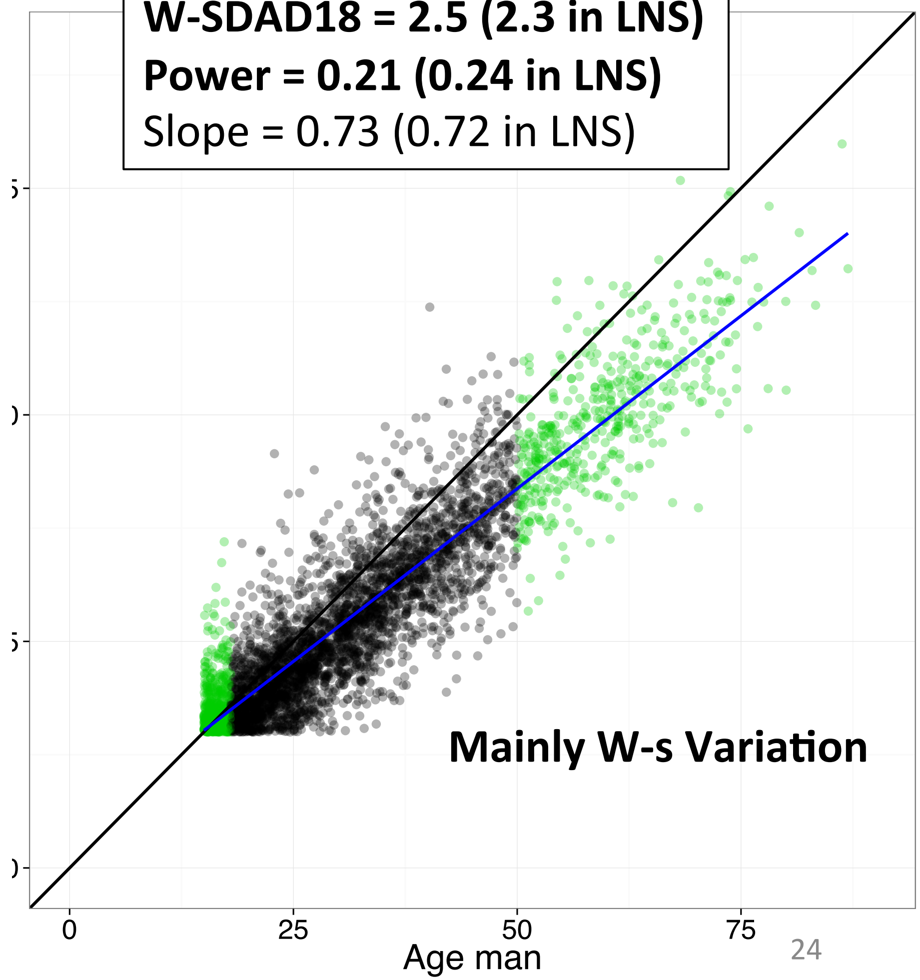


Alternative models

AAD = 4.8 (4.0 in LNS)
SDAD = 5.1 (4.9 in LNS)
B-SDAD = 3.6 (1.5 in LNS)
W-SDAD18 = 1.9 (2.3 in LNS)
Power = 0.14 (0.24 in LNS)
Slope = 0.77 (0.72 in LNS)

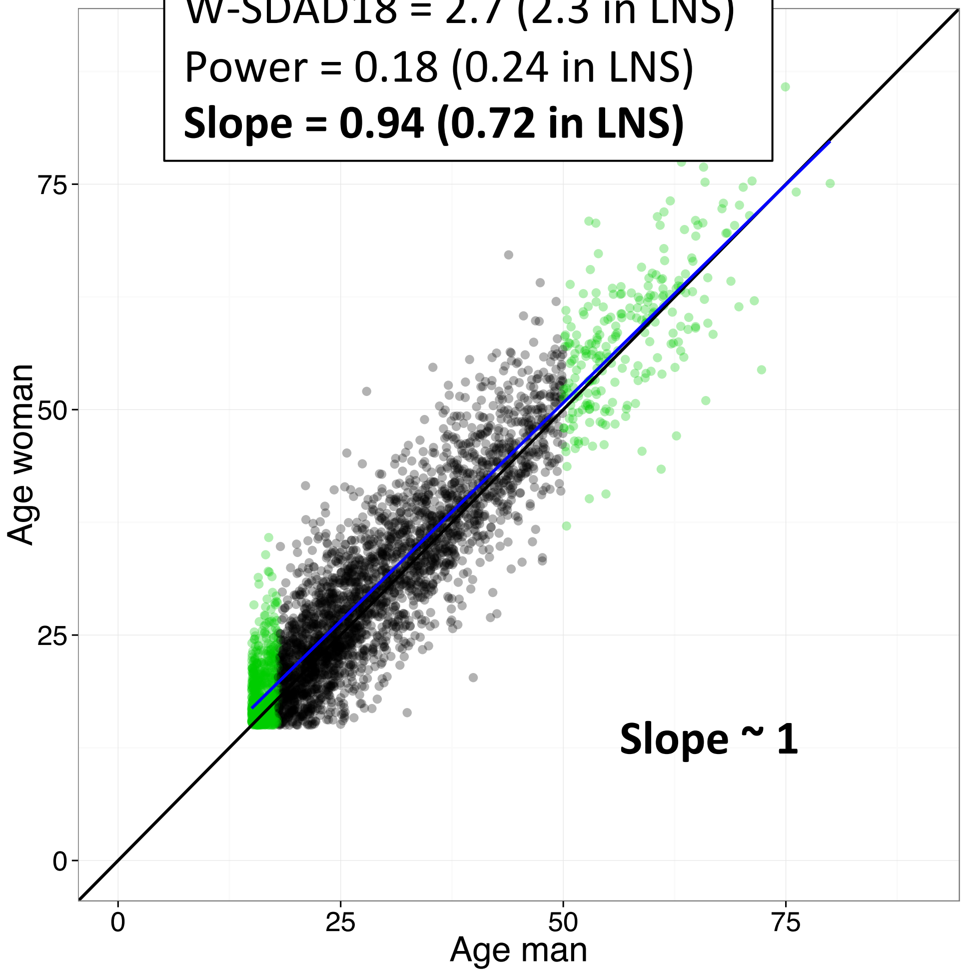


AAD = 3.9 (4.0 in LNS)
SDAD = 5.6 (4.9 in LNS)
B-SDAD = 0.8 (1.5 in LNS)
W-SDAD18 = 2.5 (2.3 in LNS)
Power = 0.21 (0.24 in LNS)
Slope = 0.73 (0.72 in LNS)

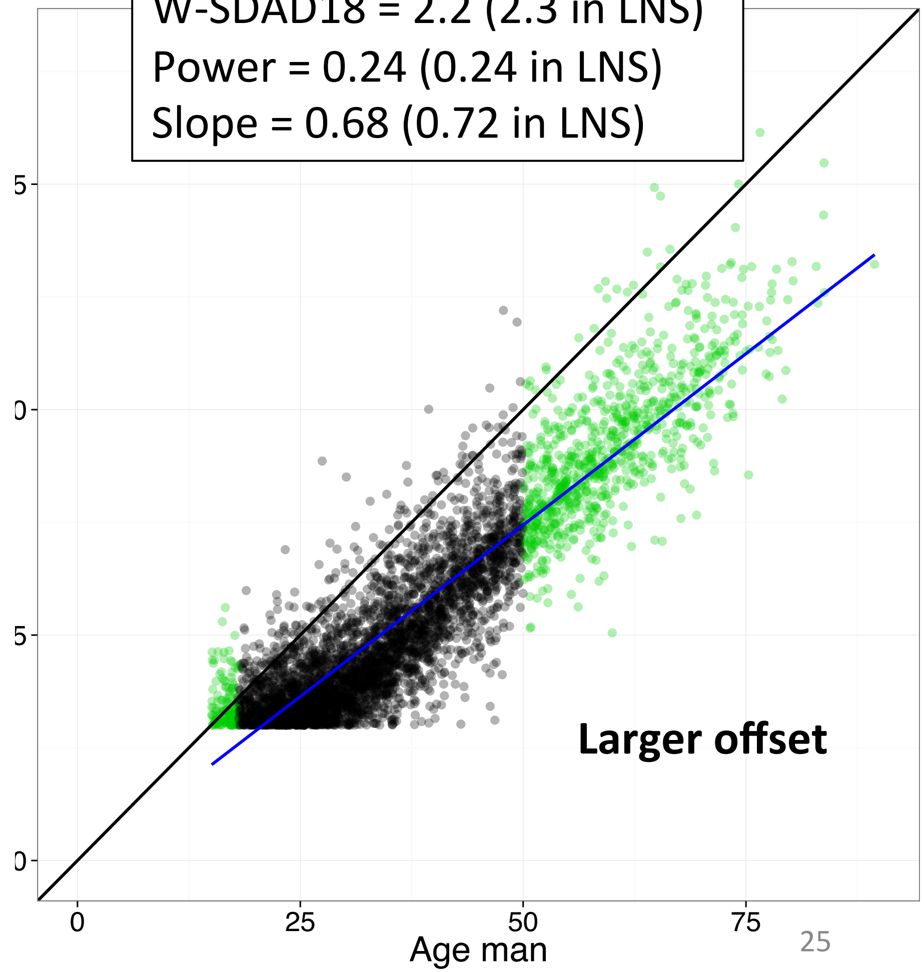


Alternative models

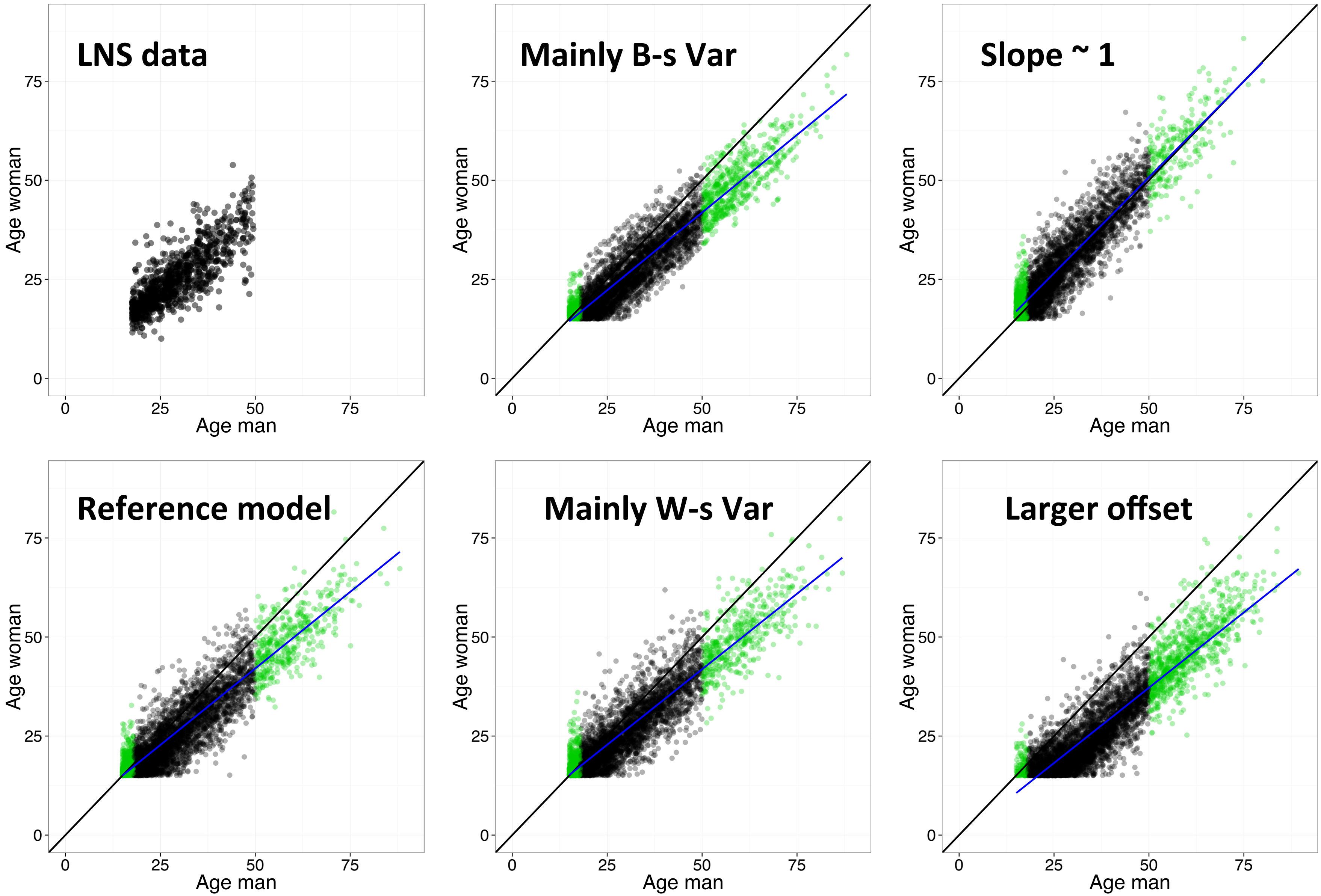
AAD = -1.4 (4.0 in LNS)
SDAD = 4.7 (4.9 in LNS)
B-SDAD = 2.1 (1.5 in LNS)
W-SDAD18 = 2.7 (2.3 in LNS)
Power = 0.18 (0.24 in LNS)
Slope = 0.94 (0.72 in LNS)



AAD = 9.7 (4.0 in LNS)
SDAD = 5.8 (4.9 in LNS)
B-SDAD = 2.1 (1.5 in LNS)
W-SDAD18 = 2.2 (2.3 in LNS)
Power = 0.24 (0.24 in LNS)
Slope = 0.68 (0.72 in LNS)



Model recap



LNS data
Mainly B-s Var
Slope ~ 1
Reference model
Mainly W-s Var
Larger offset
Age woman
Age man
0
25
50
75

Challenges

- Fitting to model to many pieces of data, but:
  - **Long** run-times
  - A **large** parameter space
  - Summary statistics from different sources

Active Learning

1. Design of Experiments

	x_1	x_2	…	x_n
Run 1	0.5	2.4	…	100
Run 2	0.5	2.4	…	400
Run 3	0.7	0.1	…	100
…	…	…	…	…
10 000	0.4	1.7	…	500

2. Simulation Model

input

response

3. Surrogate Modelling

4. System Understanding

(Willem, Stijven, et al. 2014 PLoS Comput. Biol.)

Active learning & SIMPACT

- Dataset
  - 12 input parameters
  - 8 output variables
  - Mean squared error from “targets”
- Goal
  - Understand model behavior & improve model fit
- Ongoing research
  - Total mean squared error is hard to predict
  - Predict **output variables** separately
    - Feature selection => identify driving parameters & reduce dimensionality?
    - Can we identify parameter ranges with a **prediction uncertainty**?
      - Adapt design of experiments
  - Predict **mean squared error** for each output variable
    - Feature selection => identify driving parameters?
    - Can we identify parameter ranges with a **high predicted error**?
      - Adapt design of experiments

Active learning & SIMPACT

- E.g., **Exponential growth rate**
  - Variable presence in surrogate models

    => Measure for variable importance

Variable Presence Table

Sample more into detail

Out[332]=

	# Models	% of Models	Variable	Meaning
1	276	100.0	x5	cfg_person_eagerness_dist_gamma_b
2	276	100.0	x4	cfg_person_eagerness_dist_gamma_a
3	276	100.0	x3	cfg_formation_hazard_agegapry_numrel_diff
4	184	66.7	x9	cfg_conception_alpha_base
5	155	56.2	x11	cfg_dissolution_alpha_4
6	100	36.2	x2	cfg_formation_hazard_agegapry_numrel_gender
7	46	16.7	x6	cfg_formation_hazard_agegapry_eagerness_diff
8	29	10.5	x10	cfg_conception_alpha_agewoman
9	25	9.1	x12	cfg_person_agegap_gender_dist_normal_mu
10	15	5.4	x7	cfg_formation_hazard_agegapry_gap_factor_gender_exp
11	15	5.4	x1	cfg_person_agegap_gender_dist_normal_sigma
12	2	0.7	x8	cfg_formation_hazard_agegapry_gap_factor_gender_age

Can be removed, **at least for the exponential growth rate (!)**

Active learning & SIMPACT

- E.g., **Exponential growth rate**

  Response prediction plot:

  => predicted input-output behavior if all other parameters remain constant