Community viral load

We explore how sexual behavior and STI co-infection (HSV2, HCV) dynamics jointly influence the evolution of HIV viral load, and subsequently the rate of new HIV infections.
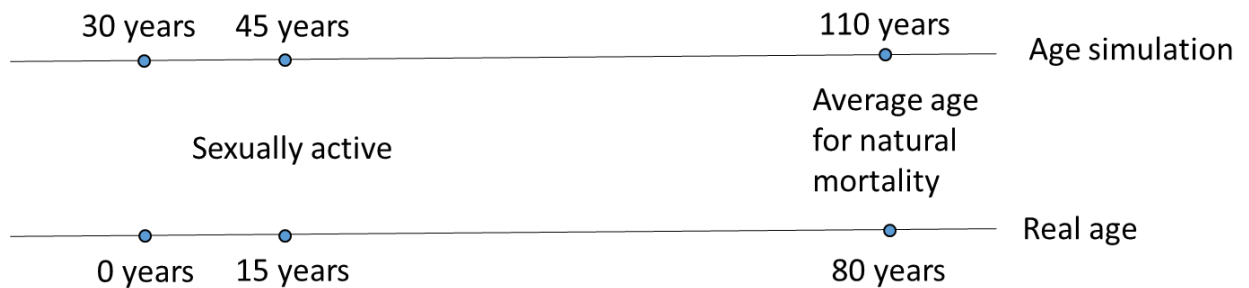
1. Implement HIV-HSV2 (resp. HCV) co-infection in SIMPACT
    Parameters HIV-HSV2: take parameters corresponding with risk ratio in the literature.
2. Simulate
    ~ 10,000 scenarios in a homosexual population of 10,000 men:
    -   introducing ART in 1994 (simulate as an intervention event, adapt every 1-2 year, vary monitoring.cd4.threshold);
    -   stabilization period = 10 years
    -   for the period 1980 – 2015 (use population pyramid Belgium 1980 to calculate parameters for the Weibull distribution)
    -   and varying the following parameters:
        1) parameters related to sexual behavior (agegap formation hazard, MSM version):
            -   $\alpha_{numrel,sum}$: penalty per ongoing relationship
            -   $\alpha_{numrel,diff}$: penalty for differing degrees (very negative value results in more assortative mixing w.r.t. degree)
            -   $\alpha_{eagerness,diff}$: penalty for differing intrinsic sexual drive (very negative value results in more assortative mixing w.r.t. intrinsic sexual drive)
            -   scale parameters of the distribution from which individuals draw intrinsic sexual drive value (E_{man1} and E_{man2}). This scale parameters influences the amount of heterogeneity in intrinsic sexual drive levels in the population. Simulate different types of distributions.
        2) parameters related to intra-couple HIV transmission
            -   rho: correlation between SPVLs of infector and newly infected individual (to study SPVL heritability)
            -   d: penalty on transmission hazard per ongoing relationship
            -   mean and stdev of the distribution from which SPVL values are drawn for individuals who acquire infection in the course of a simulation run
            -   mean and stdev from which SPVL values are drawn (investigate role of founder effects).

    We can sample using Latin Hypercube Sampling. Software: R – lhs package, function randomLHS.

    Key assumptions:

    -   heritability of set point viral load
    -   agegap formation hazard

    Because we simulate a population of only men, where no birth events can happen, we have to avoid in the simulation that a lot mortality events caused by AIDS happen. We can implement this like indicated in the figure below.

30 years   45 years   110 years   Age simulation

Sexually active   Average age for natural mortality

0 years   15 years   80 years   Real age

3. For each of the 10,000 scenarios, we calculate the following summary statistics:
    1) To characterize sexual network connectivity
       - point prevalence of concurrency at t15 (and other time points)
       - summary statistics of network connectivity:

         - degree distribution:

           For a sexual network, vertices represent persons and edges sexual

           relationships between them. The degree d of a vertex represents the

           number of sexual partners of a person, and the degree distribution is the

           distribution of d.

         - closeness centrality: high closeness centrality vertices correspond to
         highly influential individuals. People with high closeness centrality in a
         sexual network are highly responsible for the spread of HIV.

         - betweenness centrality: indicates the extent to which individuals
         transmit HIV to others within the sexual network.

         - average distances between distinct vertices: can be used to describe

           geographical distances between persons.

       - summary statistics of network assortativity:

         - assortativity coefficient: minimum (between -1 and 1) = perfectly
         dissortative, every edge connects vertices (individuals) of two different
         categories; 1 = perfectly assortitative, edges connect only vertices
         (individuals) from the same category; 0 = random.

         - for continuous characteristics (e.g. age): Pearson correlation coefficient.

       - summary statistics of network heterogeneity:

         - degree distribution

    2) To characterize time-trend in viral load distribution
       - mean and median log10 SPVL of seed HIV infections

- difference t0 – t15 in mean, median, stdev and IQR log10 viral load
- difference t0 – t15 in mean, median, stdev and IQR log10 SPVL
- also other time points than t15

3) To characterize the course of the HIV epidemic
- HIV prevalence at t15
- Cumulative HIV incidence
- Average age among HIV infected people at t15
- Also other time points than t15

4) To characterize transmission dynamics
- Average number of new infections over the complete infectious period
- Average number of new infections over the complete infectious period, per log10 SPVL bin

5) To characterize HIV-HSV2 (resp. HIV-HCV) co-infection
- HSV2 (resp. HCV) prevalence at t 15
- HIV-HSV2 (resp. HIV-HCV) prevalence at t15
- Cumulative HSV2 (resp. HCV) incidence
- Cumulative HIV-HSV2 (resp. HIV-HCV) incidence
- Average age among HSV2 (resp. HCV) infected people
- Average age among HIV-HSV2 (resp. HIV-HCV) infected people
- Also other time points than t15

Make plots of summary statistics with RSimpactHelp. If it is necessary to write new code, add to RSimpactHelp package.

4. Symbolic regression models (surrogate models) will be fitted to evaluate links between the summary statistics above. Software: Mathematica, UA (code developed by Lander Willem & Sean Stijven); R: rgp package + eventually multivator package (R).

5. Sensitivity analysis:
   - examine the role of founder effects (extremely high / low VL in seeding HIV infections) in driving the time-trend in community VL distribution;
   - study the effect of relaxing the assumption that intra-couple transmission rate is independent of the number of ongoing relationships individuals have.
     Therefore, consider the following two models
     1) intra-couple transmission is a function of the viral load
        $h = \exp(a + b*V^{\wedge}(-c))$ where V is the viral load.
     2) Sexual activity within relationships may decrease with growing numbers of concurrent relationships that man 1 (P_i) and man 2 (P_j) have.
        $h = \exp(a + b*V^{\wedge}(-c) + d*(P\_i + P\_j))$

Appendix: LHS and symbolic regression

1. Determine parameters + range
2. Design matrix (package LHS in R, 10 000 scenarios)
3. Run SIMPACT for 10 000 scenarios + postprocessing (calculate summary statistics)
4. Input-output csv (see example mail Lander)
5. UAntwerpen (Lander): symbolic regression
   - Variable importance
   - System dynamics
   - Selection of best model
1. Parameter selection + range
2. …