D4.1 Validated Individual Node Dynamics

Project acronym: *DYNANETS*
Project full title: Computing Real-World Phenomena with Dynamically Changing Complex Networks.
Grant agreement no.: 233847

Due-Date:	12
Delivery:	12
Lead Partner:	UvA
Dissemination Level:	Public
Status:	Final
Approved:	Q Board, Project Steering Group
Version:	Final

DOCUMENT INFO

0.1 Authors

Authors	Institution	e-mail
Rick Quax	UvA	r.quax@uva.nl

0.2 Document History

Date	Version	Editor	Change	Status
29-05-2010	1.0	OpenOffice		Draft
14-06-2010	1.1	Word		Draft

0.3 Glossary

	Description
MSM:	Men who have sex with men
IDU:	Injecting drug user
CN:	Complex network
HIV:	Human Immunodeficiency Virus
AIDS:	Acquired Immuno-Deficiency Syndrome

TABLE OF CONTENTS

TABLE OF FIGURES

1 Introduction

Despite the availability of a large number of mathematical models describing the spreading of epidemics such as HIV, a good understanding of the spreading dynamics through numerical analyses is still a major challenge. Problems arise from the various routes of infection, behavioral changes due to infection status, treatment effects, and inhomogeneity of the involved population and their interactions. It is essential to combine epidemiological processes with sociological models and network sciences, yet traditional modeling methodology lacks this.

In order to address these problems we combine complex networks, which can model various interaction types and routes of infection, with agent-based modeling, which enriches the network's nodes with individual node dynamics. These node dynamics enable modeling crucial aspects such as behavioral changes, treatment effects, and inhomogeneity of the involved population. This is needed to adequately model the spreading of epidemics such as HIV, which is characterized by these problems.

As a prerequisite to designing the individual node dynamics for our own simulations, in this deliverable we first explore the validation of specific sets of individual node dynamics. We present three studies conducted by members of WP4 on the design, simulation and validation of specific sets of individual node dynamics for HIV and influenza. These *use-cases,* combined with the validated network structures from WP3, will provide the foundation for the model development in DynaNets.

2 Use Cases

2.1 Stochastic simulation of HIV population dynamics through complex network modeling.

In this study we start with the formalization of 'dynamics' on and of networks, forming the foundation of our methodology. It builds a bridge between the classic mathematical modeling (SIR) and individual-based simulation on complex networks. Each process of node dynamics is regarded as a 'dynamics operator' Γ_i subscript which stochastically maps an instance of a complex network <V,E> to a new version. All dynamics operators combine together as Γ to progress the network one time step ahead, as follows.

$$\Gamma \equiv \otimes_k \Gamma_k \qquad (1)$$

$$\begin{aligned} \langle V, E \rangle_{t+1} &= \langle V, E \rangle \\ \langle V, E \rangle_{t=0} &\equiv \langle V_0, E_0 \rangle \end{aligned} \qquad (2)$$

We present a parameterized CN model describing the dynamics of HIV spreading which takes into consideration all the existing kinds of HIV spreading. Homosexual and heterosexual spreading is described by a scale-free network, drug users spreading is described with the assumption of homogeneous mixing inside the exposure group. All the network parameters have been taken from the medical literature and were fixed during the numerical experiments.

Node dynamics

The node dynamics are as follows:

1. A susceptible node can become infected with HIV by a neighbor;
2. A HIV-infected node can progress into AIDS given time;
3. A HIV-infected node can be treated, which reduces its infectivity. The coefficients for different treatment epochs (pre-ARV, ARV, and HAART-treatment) are included and used in their specific time spans;
4. A node can be removed from the network;
5. A node can be added to the network.

Complex Network

The network structure is taken to have a power-law degree distribution with exponent 1.6. In each time step (year), the network is newly generated. Therefore the network is purely random over time; there is no correlation between the existence of previous edges and the formation new edges.

Validation

All the network parameters have been taken from the medical literature; infectiousness, node removal and node addition rates have been fitted to historic data by means of least-squares estimation using a standard method steepest descent with adaptive step sizes. The historic data is the official data and estimation by the Center for Disease Control [CDC07], which

represents approximately 40,000 new transmissions per year all the way from 1990 to the present time.

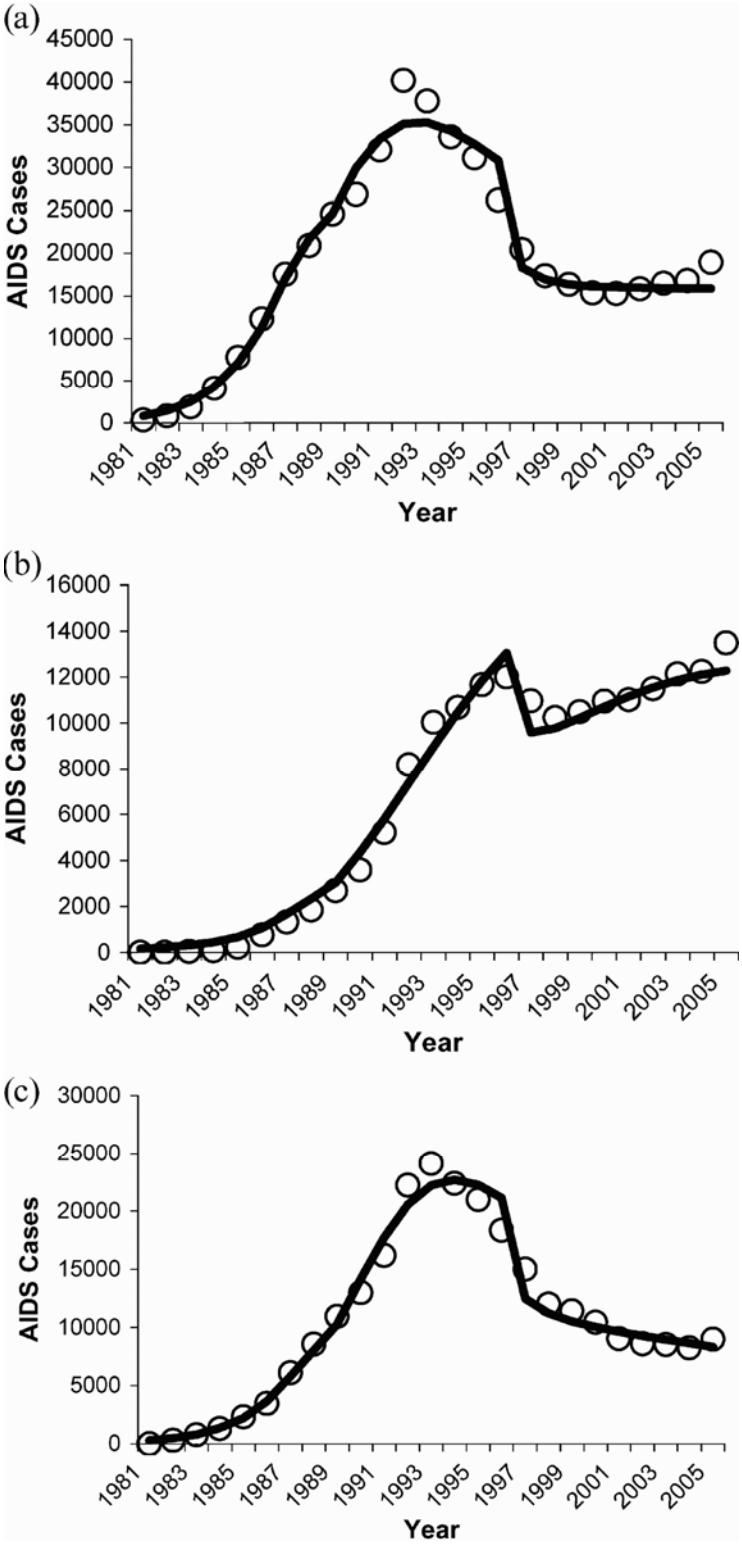

Figure 1: Simulation results (line) and reported data (circles) for the AIDS epidemic in USA of the (a) MSM exposure groups, power-law distribution with exponent γ=1.6 and k_{max}=250, I_0=0.32%, d=0.04, λ=0.44; (b) Heterosexual exposure group, power-law

distribution with γ=2.7 and k_{max}=60, I_0=0.2%, d=0.05, λ=0.28; (c) IDU exposure group, homogeneous mixing, simulation parameters are I_0=0.16%, d=0.025, λ=0.72. (I_0 is the initial epidemic size; d is the fraction of nodes that are added and removed; γ is the exponent of the power-law degree distribution; and λ is the per-link probability of infection (infectiousness).) Circles are the annual officially registered number of AIDS cases; the solid line indicates the simulation results.

As shown by the diagram above, the simulation results and the officially registered number of AIDS cases in the USA correspond remarkably well. Naturally, one should consider these result very carefully, since some of model deficiencies may have been diminished by the free 'tunable' parameters. However, one can clearly distinguish the effect of HAART treatment which are accurately captured by the simulation. Also, the stabilization of the epidemic after initial harmonic fluctuation is captured.

Parameter values

Please refer to Reference [SLO08] for details.

2.2 Modeling and simulating the propagation of infectious diseases using complex networks.

Here we take the formalization of the previous study and expand the set of individual node dynamics, focusing on the MSM population. Also, all model parameters are taken from literature so no parameter fitting takes place; this strengthens the claim that any correspondence to historic data must be the result of both the individual node dynamics and the network structure.

The main goal of the simulation was to show that key HIV dynamics by themselves could already result in complex incidence and prevalence statistics, without explaining increases or decreases using hypothetical causes such as introductions of treatments, changes of individual behavior, or migration of infected persons from elsewhere. For this reason we have implemented only one treatment type (with ARV-like coefficients) which started to be available from time zero, so that any observed characteristics of incidence and prevalence plots are simply due to node dynamics and network structure.

Node dynamics

1. A susceptible node can become infected with Acute HIV by a neighbor;
2. A node wih Acute HIV progresses to Asymptomatic HIV after three months (and has much higher infectiousness);

3. A node with Asymptomatic HIV can progress to AIDS given time, depending on whether it receives treatment;
4. A node may be diagnosed with HIV, reducing its risky behavior probability;
5. A HIV-infected node can be treated, which reduces its infectiousness and increases the time duration of the Asymptomatic HIV period. The coefficients for only one treatment type (ARV) is included and kept constant throughout the simulation;
6. A node can fail treatment;
7. A node can be removed from the network;
8. A node can be added to the network (coincides with node removal for constant population size);
9. A node can add an edge (relationship): choosing the neighboring node is dictated by a hierarchical description of edge forming probabilities;
10. A node can remove an edge (relationship);

Complex network

The network structure is taken to have a power-law degree distribution with exponent 1.6. In addition, we impose a statistically regular community structure using the Kronecker algorithm and fitting its four parameters to the degree distribution, maximum degree and network size. In short, this algorithm creates two sub communities of nodes with distinct edge probabilities and recursively divides the sub communities until it reaches communities of size 1.

The network is generated once for bootstrapping, but in subsequent time steps only the node dynamics dictate which edges are removed or added: no global network generation procedure is invoked in any time step unlike in the previous study. Thus there exists correlation between the presence of previous edges and the formation of new edges.

Validation

All the parameters of individual node dynamics have been taken from the literature. The historic data (as collected in Reference KAT02) are official data and estimates of AIDS incidence, AIDS deaths, AIDS prevalence, and HAART use, which were derived from the San Francisco Department of Public Health AIDS registry (Illustration 2). Because of the high uncertainty in HIV incidence, we use prevalence statistics mainly of AIDS for comparison, which are the most reliable data.

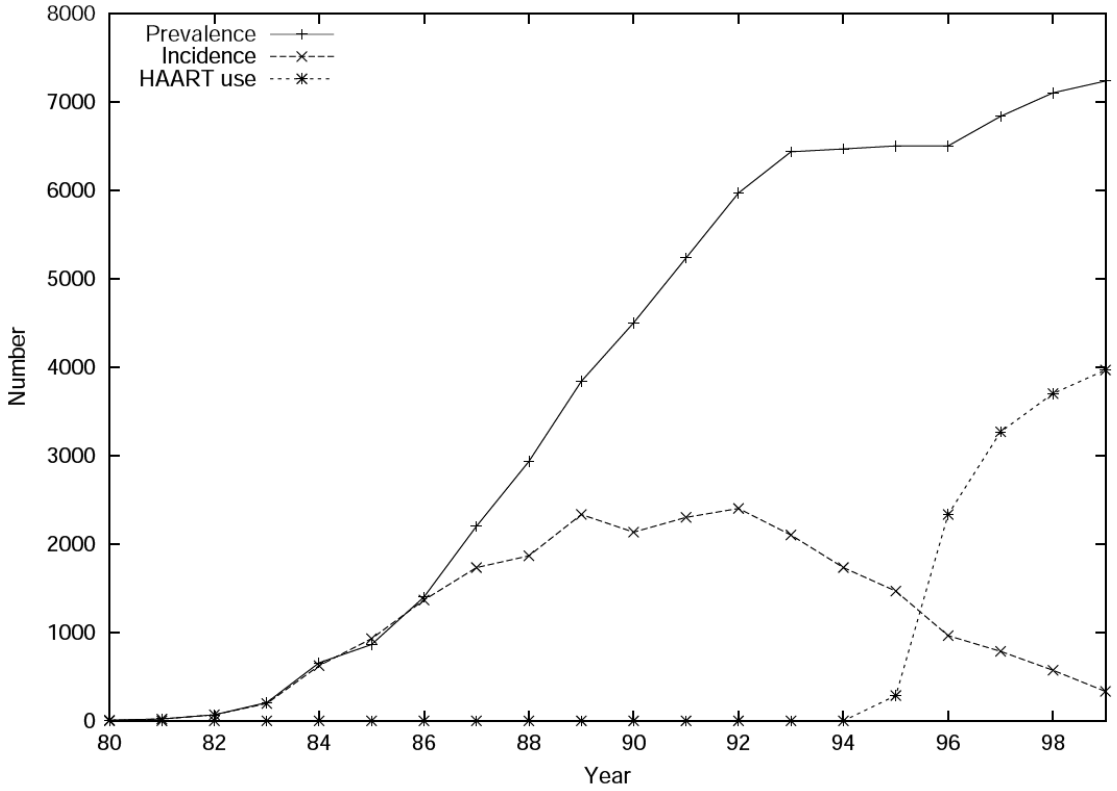

Figure 2: Trends of AIDS incidence and prevalence in San Francisco, and the use of HAART.

The simulation results for prevalence are as follows, averaged over four runs. The start and end time of the simulation and historic data (as collected and analyzed in References KAT02 and SCH01)are the same, so the horizontal scales are equal.

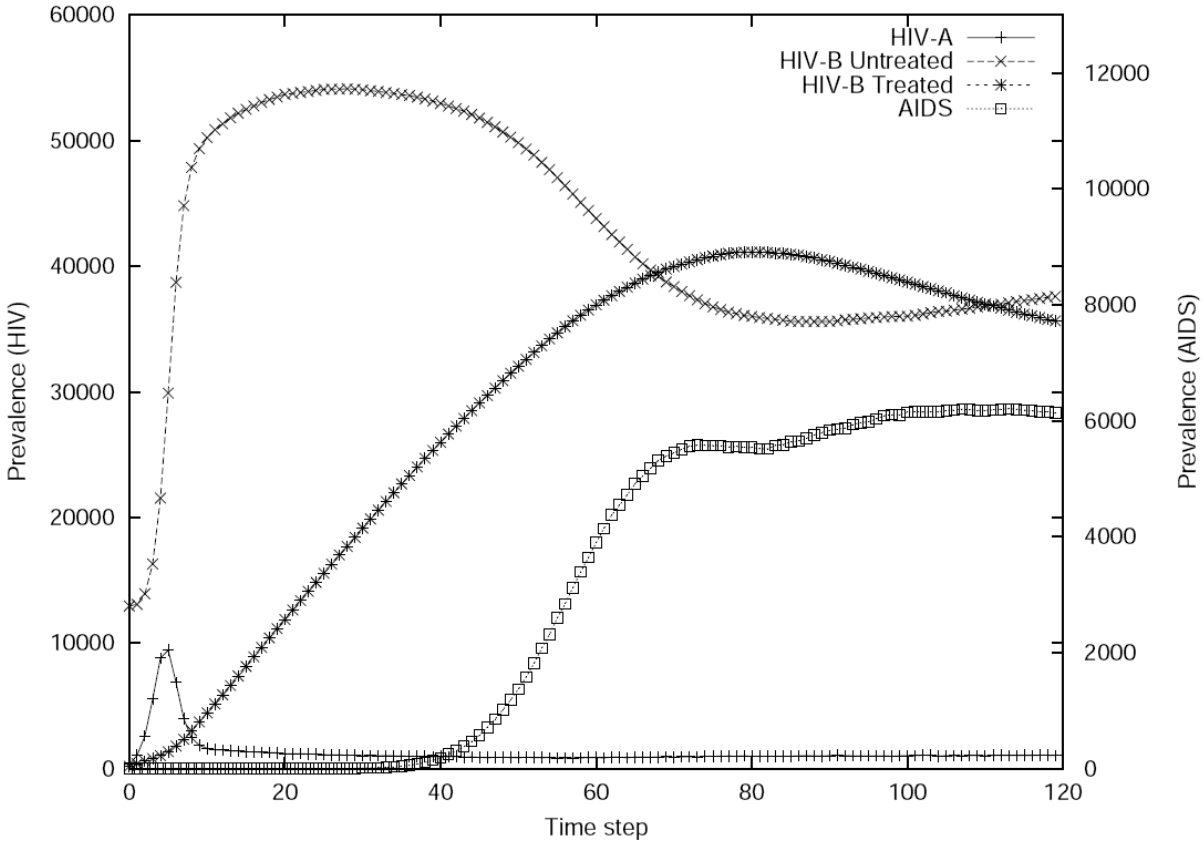

Figure 3: Simulation results of the prevalence of various stages of HIV, averaged over four runs. Each time step corresponds to three months. HIV-A is the undiagnosed accute phase which lasts for about three months; HIV-B is the asymptomatic period.

Of particular note is the diagram for AIDS prevalence, which are the squares in Illustration 4 and pluses in Illustration 3. Although no parameter is fitted to the data, the time around which the prevalence starts to take up, levels off for the first time, 'bumps' up, and levels off again correspond remarkably well with the historic data.

Another notable observation made in the study is that prevalence of untreated HIV in the simulations (crosses) peaks at around time step 25, which corresponds to year 1986, while reports [e.g. BRO91] estimate the peak of HIV incidence must have been around 1985.

Parameter values

Please refer to Reference [QUA08] for details.

2.3 Complex agent networks explaining the HIV epidemic among homosexual men in Amsterdam.

This study slightly expands the set of individual node dynamics of the previous study and applies it to Amsterdam. This means that the coefficients used in the model and the historic data for comparison are that of Amsterdam, and that the population is relatively small. Methods that rely on big numbers in order to be statistically correct would not work in this case. To address this small population size we have developed Complex Agent Network, which is the embedding of agent-based modeling in complex networks.

Node dynamics

1. A susceptible node can become infected with Acute HIV by a neighbor, where the probability depends on condom use, nature of relationship (steady or casual) and risky behavior;
2. A node with Acute HIV progresses to Asymptomatic HIV after three months (and has much higher infectiousness);
3. A node with Asymptomatic HIV can progress to AIDS given time, depending on whether it receives treatment;
4. A node may be diagnosed with HIV, reducing its risky behavior probability;
5. A node can have variant risk behavior, with coefficients from literature that identifies ranges of years per coefficient value;
6. A HIV-infected node can be treated, which reduces its infectiousness and increases the time duration of the Asymptomatic HIV period. The coefficients for only one treatment type (ARV) is included and kept constant throughout the simulation;
7. A node can fail treatment;
8. A node can be removed from the network;
9. A node can be added to the network (coincides with node removal for constant population size);
10. A node can add an edge (relationship). An edge can be one of two types: a steady relationship, which lasts for some extended period of time and has many sexual interactions per time unit, and casual, which is incidental;
11. A node can remove an edge (relationship);

Complex Network

The network is firstly generated as a random network with power-law degree distribution with exponent 1.6. Then the node dynamics dictate with local rules how the incident edges are added or deleted, but there is no global network generator algorithm.

Validation

We validate the results of our simulations both in terms of HIV incidence and AIDS prevalence. Please note that the AIDS prevalence statistic is only available nation-wide for the Netherlands, so in this case we can only validate the general trend.

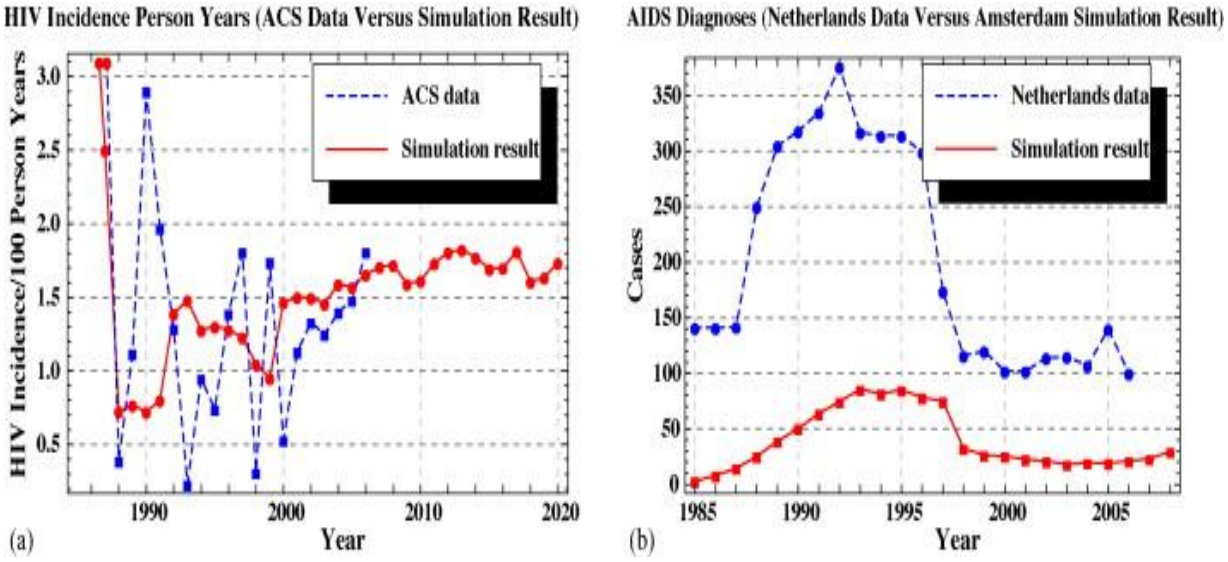

Figure 4: Comparison of the AIDS prevalence statistic between the simulation results (averaged over 24 runs) and the national AIDS prevalence data [RIVM06].

Our simulated yearly incidences are consistent with the historical ACS data [ACS06], based on a null hypothesis test in which the simulation result of incidence does not differ significantly from the ACS data: the hypothesis is accepted by using Chi-Square tests with a significance level of 0.05 (data not shown). The fluctuations in the ACS data is due to the varying and low sample sizes.

For the AIDS prevalence statistic (Illustration 4), our simulation result shows a similar trend to the statistical data in the Netherlands over the calendar years, albeit at different scales due to the difference in population sizes. Both curves show a roughly continuous increase in the number of AIDS diagnoses until 1996, then an acute decline in 1996–1998, and then a stabilization from 1999 onward.

Parameter values

Please refer to Reference [SHA09] for details.

3 Conclusion

Using our formalization of individual node dynamics coupled with network dynamics, we have studied the spreading of HIV in multiple geographical locations (United States of America, San Francisco, Amsterdam), at multiple spatial scales (continent; city), through multiple

transmission modes (IDU, heterosexual, MSM) and at various level of detail in terms of model parameterization. To this end we have incorporated a multitude of parameters from literature. In each study we find that our model's simulation results correspond remarkably well with the historic data. We conclude that our methodology of modeling individual node dynamics, coupled with network dynamics, is sufficiently powerful to model non-trivial network spreading processes such as that of HIV.

4 References

[CDC07] Bartlett, J. G. (2007) *Routine opt-out testing for HIV: rationale and obstacles* — Available at slides/transcript at http://www.medscape.com/viewarticle/551906_2

[BRO91] Brookmeyer, R., "Reconstruction and future trends of the AIDS epidemic in the United States," Science, vol. 253, no. 5015, pp. 37–42, 1991.

[SCH01] Schwarcz, S., Kellogg, T., McFarland, W., Louie, B., Kohn, R., Busch, M., Katz, M., Bolan, G., Klausner, J., and Weinstock, H., "Differences in the temporal trends of HIV seroincidence and seroprevalence among sexually transmitted disease clinic patients, 1989-1998: Application of the 65 serologic testing algorithm for recent HIV seroconversion," Am. J. Epidemiol., vol. 153, no. 10, pp. 925–934, 2001.

[RIVM06] M. van Veen, F. Koedijk, I. vanden Broek, E. op de Coul, I. de Boer, A van Sighem, M. vander Sande, Sexually transmitted infections in the Netherlands in 2006 (2006), http://rivm.openrepository.com/rivm/bitstream/10029/16495/1/210261003.pdf.

[ACS06] Overview of the Amsterdam cohort studies among homosexual men and drug users (2006). http://www.amsterdamcohortstudies.org/menu/reports/ACSoverview2006.pdf.

[QUA08] R. Quax: *Modeling and Simulating the Propagation of Infectious Diseases Using Complex Networks*, MSc thesis, Georgia Institute of Technology, Atlanta, GA, USA, (Supervisor: David Bader) August 2008.

[SHA09] S. Mei; P.M.A. Sloot; R. Quax; Y. Zhu and W. Wang: *Complex Agent Networks explaining the HIV epidemic among homosexual men in Amsterdam*, Mathematics and Computers in Simulation, vol. 80, nr 5 pp. 1018-1030. Jan 2010. (DOI: 10.1016/j.matcom.2009.12.008)

[SLO08] Sloot, P. M. A., Ivanov, S. V., Boukhanovsky, A. V., Van De Vijver, D. A. M. C. & Boucher, C. A. B. (2008). Stochastic simulation of HIV population dynamics through complex network modelling. *International Journal of Computer Mathematics*, *85*(8), 1175-1187. doi:10.1080/00207160701750583.

[KAT02] Katz, M. H., Schwarcz, S. K., Kellogg, T. A., Klausner, J. D., Dilley, J. W., Gibson, S., and McFarland, W., "Impact of highly active antiretroviral treatment on HIV seroincidence among men who have sex with men: San Francisco," Am J Public Health, vol. 92, no. 3, pp. 388–394, 2002.