Appendix 'A'
Outline of the individual-based episodic risk model

We describe the stochastic individual-based (IBM) version of the deterministic compartmental model (DCM) of episodic risk described in Zhang et al. (2012). Table A-1 gives a list of the model parameters and Table A-2 gives the set of differential equations from the DCM. This model was used to generate acute infection outbreak clusters, which are reported in the main text. Notice that the recorders and summary generators for outbreak clusters can be used with any individual-based model of HIV transmission and are described in Appendix 'B'.

Table A-1: List of the episodic risk model parameters adapted from Zhang et al. (2012)

Parameter	Value	Unit	Definition
χ	3.4114	/month	Average contact rate in the entire population
μ	1/(40*12)	/month	Flow of new individuals into sexually active population
μ	1/(40*12)	/month	Rate of leaving the sexually active population
γ_1	1/2	/month	Rate of transitioning from acute to chronic infection
γ_2	1/120	/month	Rate of death from AIDS during chronic infection
β	0.003	/contact	Average transmission probability across stages
X	variable	-	Fraction of transmission potential from acute stage
ϕ_H *(durH)*	variable	/month	Rate of transitioning from high contact rate state to low contact rate state
ρ *(rCHL)*	variable	-	Ratio of high contact rate over low contact rate
π_H *(FrH)*	variable	-	Average fraction of population with high contact rate in the absence of HIV
υ *(FHatH)*	variable	-	Fraction of contacts of individuals with high contact rate at the high-risk site
β_1	$X\beta\left(1+\frac{\gamma_1}{\gamma_2+\mu}\right)$	/contact	Transmission probability during acute stage
β_2	$(1-X)\beta\left(1+\frac{\gamma_2+\mu}{\gamma_1}\right)$	/contact	Transmission probability during chronic stage
χ_H	$\rho\chi_L$	/month	Contact rate for the high-risk population
χ_L	$\frac{\chi}{(1-\pi_H+\rho\pi_H)}$	/month	Contact rate for the low risk population
ϕ_L	$\phi_H\left(\frac{\pi_H}{(1-\pi_H)}\right)$	/month	Rate of transitioning from low contact rate state to high contact rate state
λ_H	$\frac{A_H\beta_1+C_H\beta_2}{S_H+A_H+C_H}$	-	Force of infection per contact at the high risk mixing site
λ_L	$\frac{(A_L+(1-\upsilon)A_H\rho)\beta_1+(C_L+(1-\upsilon)C_H\rho)\beta_2}{S_L+A_L+C_L+(S_H+A_H+C_H)\rho(1-\upsilon)}$	-	Force of infection per contact at the general mixing site

Table A-2: Differential equations for the deterministic episodic risk compartmental model (source: Zhang et al 2012)

$$\frac{dS_H}{dt} = U\pi_H - S_H\chi_H\left(\nu\lambda_H + (1-\nu)\lambda_L\right) - \phi_H S_H + \phi_L S_L - \mu S_H$$

$$\frac{dS_L}{dt} = U(1-\pi_H) - S_L\chi_L\lambda_L + \phi_H S_H - \phi_L S_L - \mu S_L$$

$$\frac{dA_H}{dt} = S_H\chi_H\left(\nu\lambda_H + (1-\nu)\lambda_L\right) - A_H\gamma_1 - \phi_H A_H + \phi_L A_L - \mu A_H$$

$$\frac{dC_H}{dt} = A_H\gamma_1 - C_H\gamma_2 - \phi_H C_H + \phi_L C_L - \mu C_H$$

$$\frac{dA_L}{dt} = S_L\chi_L\lambda_L - A_L\gamma_1 - \phi_L A_L + \phi_H A_H - \mu A_L$$

$$\frac{dC_L}{dt} = A_L\gamma_1 - C_L\gamma_2 - \phi_L C_L + \phi_H C_H - \mu C_L$$

In the IBM, rates of flows from the DCM introduced in Zhang et al. (2012) were converted to probabilities applied to the simulated individuals (see Table A-1). For a given rate of flow (λ), a probability of transition is calculated as $1-e^{-\lambda\, dt}$,
The simulation schedule in the individual based model consists of three main processes called at each time step. As in the DCM, initial HIV infection is introduced to 1% (default) of the total population. We assume instantaneous sex acts among the individuals.

A list of all possible events at the individual level is given as follows:

1. An individual in high-risk phase selects another individual in high-risk phase at the high-risk mixing site.
2. An individual in high-risk phase selects another individual in high-risk phase at the common mixing site.
3. An individual in low-risk phase selects another individual in high-risk phase at the common mixing site.
4. An individual in low-risk phase selects another individual in low-risk phase at the common mixing site.
5. An acute stage individual in high-risk phase infects a susceptible individual in high-risk phase at the high-risk mixing site.
6. An acute stage individual in high-risk phase infects a susceptible individual in high-risk phase at the common mixing site.
7. An acute stage individual in low-risk phase infects a susceptible individual in high-risk phase at the common mixing site.
8. An acute stage individual in low-risk phase infects a susceptible individual in low-risk phase at the common mixing site.
9. A chronic stage individual in high-risk phase infects a susceptible individual in high-risk phase at the high-risk mixing site.
10. A chronic stage individual in high-risk phase infects a susceptible individual in high-risk phase at the common mixing site.
11. A chronic stage individual in low-risk phase infects a susceptible individual in high-risk phase at the common mixing site.

12. A chronic infected individual in low-risk phase infects a susceptible individual in low-risk phase at the common mixing site.
13. An individual in the low-risk phase switches to the high-risk phase.
14. An individual in the high-risk phase switches to the low-risk phase.
15. An individual in the acute infection stage progresses to the chronic stage.
16. An individual in the chronic infection stage dies of AIDS.
17. A new susceptible individual enters the system.
18. An existing individual leaves the system with an average duration of stay set as 40 years.

The simulation schedule is given below:

```
For i ← 1 to maximum-iterations
        Call Update-Individuals-Status
        Call Risk-Phase-Transitions
        Call Sexual-Interaction-of-Individuals
End
```

Update-Agents

At each time step, individuals leave the system with an exit rate μ (see Table A-1). Susceptible individuals are introduced into the system, with the same rate μ. Infected individuals status of infection is also updated. An infected individual advances from the acute to chronic stage based on the probability calculated from the rate β_1. Progression from the chronic stage to dying of AIDS is determined by the probability from the rate β_2.

Risk-Phase-Turnover

Individuals in the IBM switch their risk phases (from high to low and vice versa) based on the probabilities that are determined by the risk turnover parameter π_H, which is the average duration of stay of in the high-risk (high contact rate) phase for an individual. Susceptible individuals entering in the system during the simulation are assigned risk groups depending upon the e quilibrium frequency of the two risk groups.

Sexual-Interaction-of-Individuals

In the IBM, the proportion of individuals in high-risk phase is determined by the model parameter π_H (see Table A-1). High-risk individuals select other individuals based on their preference for assortative mixing. This is determined by the model parameter υ, which is the fraction of high-risk contacts that are made at the high-risk mixing site; the rest of the partners are selected from the common mixing site. Individuals in the low-risk phase have sexual interaction with other individuals at the common mixi ng site only.

References

Zhang, X., Zhong, L., Romero-Severson, E., Alam, S.J., Henry, C., Volz, E.M., Koopman, J.S., 2012. Episodic HIV Risk Behavior Can Greatly Amplify HIV Prevalence and the Fraction of Transmissions from Acute HIV Infection. Statistical Communications in Infectious Diseases.

Appendix 'B'
Technical Description of the Acute Infection Outbreak Java Library

1. Introduction

This library is currently being developed in Java 1.6. The source code along with the documentation is available at the HIV Risk Dynamics project's Google code repository: http://code.google.com/p/hiv-risk-dynamics/ under the Creative Commons 3.0 license. This is a work in progress. For an updated technical description and source code, the reader is referred to the afore-mentioned link.

1.1 Setting up the code

The project available at the Google code repository can be directly imported in the Eclipse IDE (or other IDEs). The code must be linked to the required additional open source Java libraries that are listed below.

Additional libraries required:

- JUNG (Java Universal Network/Graph Framework) version 2.x
  http://jung.sourceforge.net/
- Michael Thomas Flanagan's Java Scientific Library
  http://www.ee.ucl.ac.uk/~mflanaga/java/
- Apache's Math Common – http://commons.apache.org/math/
- CERN's Colt Library – http://acs.lbl.gov/software/colt/

The library is composed of two parts: generation of acute infection outbreaks data and the summarizing population-level outputs and measures of the generated outbreaks size and shapes.

1.2 Overview of the packages

The library composes of six packages. An overview of each of the packages is given below (in alphabetical order):

- ***interfaces:*** This package contains the ParameterInterface interface that contains all constants and enum classes used by other classes in the library. All classes in the library must implement this interface. It also contains the interfaces for implementing the individual-based HIV transmission model main class, the class representing individuals in the model and the class storing data related to an HIV transmission event during a model run.
- ***cluster***: This package contains Java classes related to recording acute infection outbreaks and infection trees. It also contains Java classes for outputting all and acute infection transmissions recording during the given observed periods.
- ***display:*** This package contains classes for displaying the acute infection forest as a JSwing component. The viewer and the renderer are based on the engines provided by the JUNG 2.x library.
- ***main****:* It contains two classes Controller.java and Summarizer.java. The Controller class is responsible for running the individual-based model by invoking the model main class and providing the input parameters that are read from a given input file (default: CSV). It collects population-level outputs from the model and generates output in the output directory in the project.
- ***plfit:*** The plfit package contains the wrapper class and the original code used by the Baek et al. (2011) for estimating the distribution parameters of a data that follows a heavy-tailed

distribution. It also contains a wrapper class for accessing the C binary by Abramson et al. (2011) that fits a power-law distribution for a given data.

- ***reader:*** This package contains all the classes related to producing summaries of population-level and acute infection outbreaks related statistics that are read from the generated data. All classes in this package implement the ParametersInterface in the *base-model* package. The classes in this package are called from the Summarizer class in the *main* package.

2. Generation of acute infection outbreaks data

In this section, we use the individual-based analogue to the deterministic model of Episodic Risk described in Zhang et al. (under review) to illustrate how to integrate the library code. We first describe the interfaces, followed by implementation of these interfaces and describing classes related to the Episodic Risk model and recording of the AHI outbreaks.

2.1 Interfaces in the library

The three Java interfaces are ParameterInterface, AgentInterface and the BaseModelInterface and are placed in the *basemodel* package. For details on the methods in this interface, refer to the JavaDocs.

2.1.1 ParametersInterface: We start with the ParametersInterface, which contains all constant parameters and enum classes. This interface is implemented by all classes used in the library. Both the data generation classes and the summary outputs generation classes use this interface. The model related constants include the maximum number of iterations (simulation ticks), initial population of agents, proportion of initial infected individuals, duration of sexual activity, duration of acute and chronic stages of infection, base transmission probability etc., depending upon the features of the simulated model.

The Outbreaks related enum types are also stated in this interface:

- *OutbreakType* enum gives the option of selecting the type of an acute infection outbreak, which could be defined by the period spent in AHI (default), first six months of infected period and/or first two years of infected period.
- *OutbreakRecord* enum class defines the periods during simulation for which the acute infection outbreaks are to be recorded. Currently, there is two such periods: Acute Period and the Endemic. Users can change the start record time and the number of years for which the outbreaks to be recorded in the constructor of this enum class.

- *ChainsType* enum defines the Types of transmission chains to be considered for measuring outbreaks' statistics: continuous chains (ongoing transmission chains), dead-ends (dead transmission chains) or all chains.

- *NewChainThreshold* constant defines the threshold that allows creating of entire new transmission chains during the recording period. This means that if a person transmits infection after this threshold, we consider the newly infected person as starting a new chain as a new root of an infection tree. For details of the rest of the constants and the enum classes, refer to the JavaDocs of the library.

2.1.2 AgentsInterface: This interface provides the methods that must be implemented by a class for an individual in the individual-based model. The outbreaks recorder class ClusterRecorder in the *cluster* package and other classes in the *cluster* package use these methods to retrieve an individual agent's behavioral and infection states. Likewise, the Summarizer class and the

associated classes in the *reader* package use the setter methods in the interface to reconstruct the infection forest and compute tree statistics.

2.1.3 BaseModelInterface: This interface must be implemented by any individual-based HIV transmission model. The model assumes a fixed time step schedule and is the main engine class responsible to connect all components of the IBM (e.g. sexual mixing, updating of agents' status, and transmission events). An implementation of this interface must contain the outbreaks recorder ClusterEngine class as a member and implement the methods of this interface.

2.1.4 TransmissionInterface: The TransmnissionInterface ensures the basic information required to store a transmission event record. The toString() method outputs the transmission event to the ClusterEngine's output recorder to output all transmission and acute infection transmissions in a CSV file. Each transmission event is output as a row in the CSV file. One can extend this interface to account for more state variables about the infector and the infected agent in a more complicated model. See, e.g. the BaseTransmission class in the *cluster* package that implements this interface and the EpisodicRiskTransmission class in the *episodicriskmodel* package that extends the BaseTransmission class as an illustration.

2.2 Integrating the individual-based model of episodic risk

In this section, we illustrate how to use the library code using the Episodic Risk model (see the episodicriskmodel package). First, we implement the BaseModel interface in the EpisodicRiskModel class. This class is the main class responsible for contain collections of individuals, risk transitions, sexual mixing, birth, death, and status update of the agents. It contains the member clusterRecorder, which is the instance of the ClusterRecorder class. The clusterRecorder is created in the setup function by calling the implemented function createClusterRecorder().

The model runs until the maximum iteration from the ParametersInterface. At each time step, the function run() is called, in which all processes in the model take place. In this run() method, we call the implemented function *callClusterRecorderStep()*. To add a transmission event in the clusterRecorder, we call the addTransmissionToClusterRecord(infector, susceptible) function inside infect(), where a susceptible agent is infected by an infector agent.

At each time step, the function updateIndividuals() is called inside the run() function in the model. Inside the updateIndividuals() function, the status of individuals is updated; also, birth and death processes take place. When an individual is dead (i.e. left the system), we call updateClusterRecord() to remove the dead individual's record from the recorder.

The basic AgentInterface is extended to the EpisodicAgentInterface to account for additional states of an agent such as Risk-state and mixing-site. Likewise, other models can extend the AgentInterface for other behavioral/epidemiological states of an agent. The Person class then implements the EpisodicAgentInterface.

Finally, the BaseTransmission class in the *cluster* package implements the interface TransmissionInterface. We then extend the BaseTransmission class by the EpisodicRiskTransmission class (in the *episodicriskmodel*) package to extend the output about the transmission event such as the risk phase of the infector and the susceptible agent and the mixing site where the transmission occurred. Note that the EpisodicRiskTransmission overrides the toString from the BaseTransmission by first calling the base class' toString() function and then appending further output.

Notice that in the createClusterRecorder() method (see above), we are required to pass the 'transmission class' as the last argument to create an instance of the cluster recorder at run-time. This 'transmission class' is either the BaseTransmission class (implementing the TransmissionInterface) or an extension such as the EpisodicRiskTransmission class.

2.3 Integrating the individual-based Partnership model

In this section, we illustrate how to use the library code using the Episodic Risk model (see the partnershipmodel package). First, we implement the BaseModel interface in the PartnershipModel class. This class is the main class responsible for containing collections of individuals, risk transitions, sexual mixing, birth, death, and status update of the agents. It contains the member clusterRecorder, which is the instance of the ClusterRecorder class. The clusterRecorder is created in the setup function by calling the implemented function createClusterRecorder().

The model runs until the maximum iteration from the ParametersInterface. At each time step, the function run() is called, in which all processes in the model take place. In this run() method, we call the implemented function *callClusterRecorderStep()*. To add a transmission event in the clusterRecorder, we call the addTransmissionToClusterRecord(infector, susceptible) function inside infect(), where a susceptible agent is infected by a infector agent.

At each time step, the function updateIndividuals() is called inside the run() function in the model. Inside the updateIndividuals() function, the status of individuals is updated; also, birth and death processes take place. When an individual is dead (i.e. left the system), we call updateClusterRecord() to remove the dead individual's record from the recorder.

The basic AgentInterface is extended to the PartnershipAgentInterface to account for additional states of an agent, similar to the EpisodicRiskAgentInterface (see Section 2.2). The PartnershipAgent class then implements the PartnershipAgentInterface.

We extend the BaseTransmission class (see Section 2.2) by the PartnershipTransmission class (in the *partnershipmodel*) package to extend the output about the transmission event. Note that the PartnershipTransmission class overrides the toString from the BaseTransmission by first calling the base class' toString() function and then appending further output.

Notice that in the createClusterRecorder() method (see above), we are required to pass the 'transmission class' as the last argument to create an instance of the cluster recorder at run-time. This 'transmission class' is either the BaseTransmission class (implementing the TransmissionInterface) or an extension such as the PartnershipTransmission class.

3. Generating summaries of outputs from the simulated data

The Summarizer class in the *main* package reads the all transmissions and AHI transmissions for the existing multiple runs of a parameter set and then outputs summary statistics, which include:

3.1 Population-level variables

For each input parameter set, we record endemic prevalence and fraction of transmissions from acute HIV infections in the simulated population during the observed period. In addition to that, the library generates empirical joint distribution of AHI outbreaks with respect to size and duration categories. The

default size and duration categories may be modified or extended. Other variable and their categories such as height can also be extended in the library.

3.2 Acute Infection Outbreaks

We examine acute infection outbreaks (clusters) with respect to several aspects based on the outputs generated in a simulation run. The list is not exhaustive although it covers more aspects of outbreak distributions than in some of the previously reported simulation-based studies such as Murray (2002), Lewis et al. (2008) and Rocha et al. (2011); and phylogenetic studies such as Brenner et al. (2011).

In addition to the descriptive statistics for the summary variables (see Table B-1), we report further statistics to explore the distribution of the outbreak size distribution. For instance, the percentile measures tell us what would be the expected outbreak size (given an input parameters setting) in terms of the ranking of outbreaks, since the observed outbreak size distribution is positively skewed. The fraction of outbreak sizes indicates the density of the size distribution when the outbreaks are binned into the above categories with respect to size. For instance, a much higher proportion of 'isolates' indicate that under the given parameters setting, a big proportion of AHI outbreaks generated just a single transmission and were short-lived. On the other hand, a higher fraction of outbreaks of size greater than 10 indicates the AHI outbreaks' role in carrying on the chains of transmission. Note that the above-mentioned categories were taken to be closer to the ones used by Brenner et al. (2011). Finally, we estimated the exponent (slope) of the outbreak size distribution from the information-theoretic approach used by Baek et al. (2011). In addition to that, we get the estimates for the largest outbreak size, which is similar to the actual maximum size of the outbreaks, under a given model setting.

Table B-1: A summary of output measures that can be generated from the recorded acute infection outbreaks

Output summary	Description
Acute infection (AHI) outbreaks during the observed period	
Size of all AHI outbreaks (Average; Median; Max; Variance)	The size of an AHI outbreak is the total number of transmissions from individuals in that outbreak
Height of all AHI outbreaks (Average; Median; Max; Variance)	The height of an AHI outbreak (tree) is the maximum distance (in terms of edges) from the root to a leaf.
Internal-to-leaf ratio of all AHI outbreaks (Average; Median; Max; Variance)	(An 'Internal' is an individual that transmitted at least once during acute stage. A 'leaf' is an individual that did not transmit during acute stage). This measure refers to the shape of the outbreaks
Duration of all AHI outbreaks (Average; Median; Max; Variance)	Duration of an AHI outbreak is determined by the difference between the time (in days) of the last transmission and the first transmission from an individual in that outbreak.
Width of AHI outbreaks (Average; Median; Max; Variance)	Width of an AHI outbreak is defined by the maximum children of a node in an outbreak tree (i.e. maximum number of secondary infections by an individual in an outbreak.
Height-to-Width ratio (Average; Median; Max; Variance)	Ratio of height and width of an AHI outbreak
Empirical cumulative size distribution	
Additional measures for outbreak size distribution	
Median size	Median, or the 50th percentile of the AHI outbreak size distribution
75th Percentile	75th Percentile of the AHI outbreak size distribution
90th Percentile	90th Percentile of the AHI outbreak size distribution
99th Percentile	99th Percentile of the AHI outbreak size distribution
Fraction isolates	Fraction of outbreaks with just a single transmission (size = 1)
Fraction until 2	Fraction of outbreaks with at most 2 transmissions (size ≤ 2)
Fraction until 5	Fraction of outbreaks with size between 2 and 5 (2 < size ≤ 5)
Fraction until 10	Fraction of outbreaks with size between 5 and 10 (5 < size ≤ 10)
Fraction greater 10	Fraction of outbreaks with size greater than 10 (10 < size)

Inter-quartile mean	Average size of outbreaks taking into account of the inter-quartile range
Exponent fit of the slope	Exponent (slope) of the outbreak size distribution from Baek et al. (2011)

3.3 Patterns of Continuous transmissions

We extend our exploration of the acute infection (AHI) outbreak size distribution to their contribution to continuous chains of transmissions. We define continuous chains as the chains of transmissions that continue throughout the observed period during which acute infection outbreaks were recorded. Those chains that terminate before this period are termed dead-ends.

For all the parameter sets, we get the cumulative distribution of the chronic infection (CHI) transmissions that link AHI outbreaks. A higher proportion of a larger number of CHI transmissions linking the AHI outbreaks could indicate smaller AHI outbreaks on the transmission chains whereas a higher proportion of CHI link counts of size 1 could indicate larger AHI outbreaks on the chains.

References

Baek, S.K., Bernhardsson, S., Minnhagen, P., 2007. Zipf's law unzipped. Journal of Physics 13.

Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distributions in empirical data. SIAM Review 51(4), 661-703.

Zhang, X., Zhong, L., Romero-Severson, E., Alam, S.J., Henry, C., Volz, E.M., Koopman, J.S., 2012. Episodic HIV Risk Behavior Can Greatly Amplify HIV Prevalence and the Fraction of Transmissions from Acute HIV Infection. Statistical Communications in Infectious Diseases.

Appendix 'C'

Plots related to Section 6.2 in the main text
(High-to-low contact ratio: 1; AHI transmission potential: 0.5)

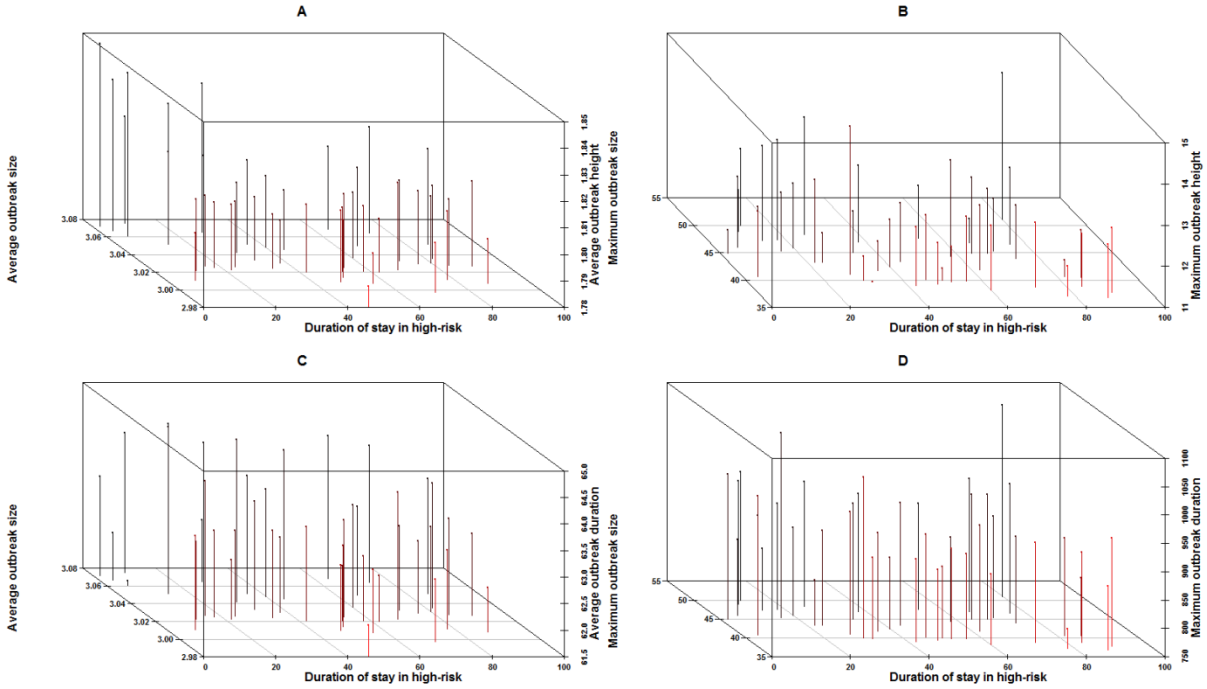

Figure C-1: 3D Scatter plots showing the relation of various AHI outbreak measures against the explored range of the 'duration of stay in high risk' for the data stratified with respect to 'high-to-low contact ratio': 1 and 'AHI transmission potential': 0.5. Panel A: Average outbreak size (Y-axis) versus average outbreak height (Z-axis). Panel B: Maximum outbreak size (Y-axis) versus maximum outbreak height (Z-axis). Panel C: Average outbreak size (Y-axis) versus Average outbreak duration (Z-axis). Panel D: Maximum outbreak size (Y-axis) versus maximum outbreak duration (Z-axis).

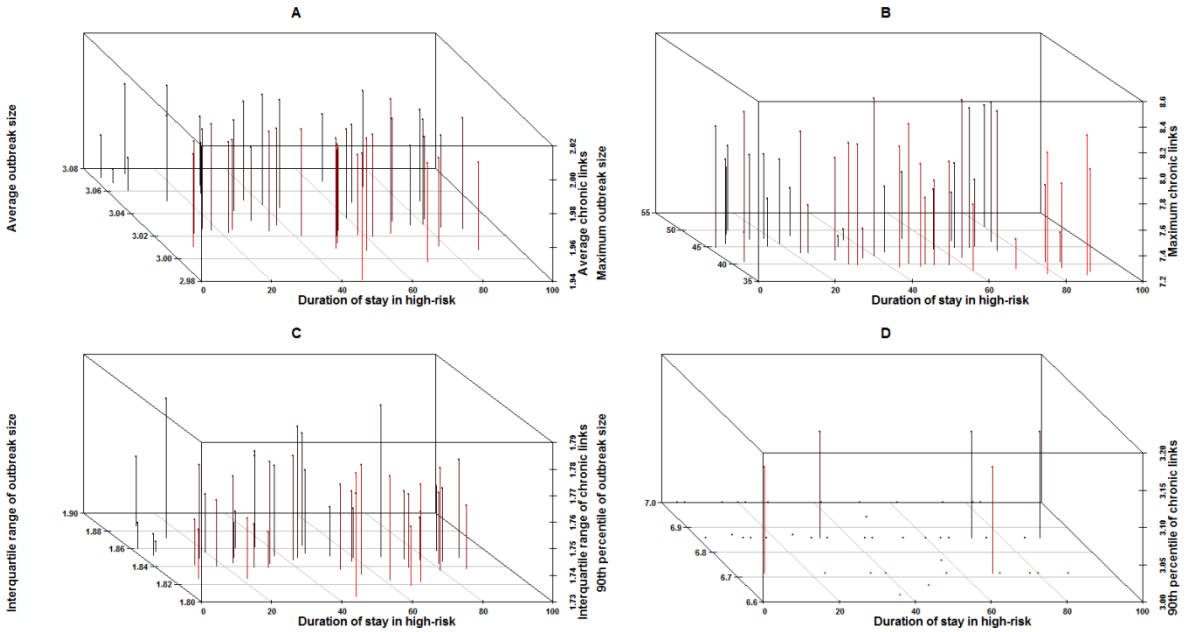

Figure C-2: 3D Scatter plots showing the relation of various AHI outbreak measures against the explored range of the 'duration of stay in high risk' for the data stratified with respect to 'high-to-low contact ratio': 1 and 'AHI transmission potential': 0.5. X-axis: Duration of stay in high-risk. Panel A: Average outbreak size (Y-axis) versus average chronic links connecting AHI outbreaks (Z-axis). Panel B: Maximum outbreak size (Y-axis) versus chronic links connecting AHI outbreaks (Z-axis). Panel C: Inter-quartile rang of outbreak size (Y-axis) versus inter-quartile range of chronic links connecting AHI outbreaks (Z-axis). Panel D: Maximum outbreak size (Y-axis) versus maximum chronic links connecting AHI outbreaks (Z-axis).

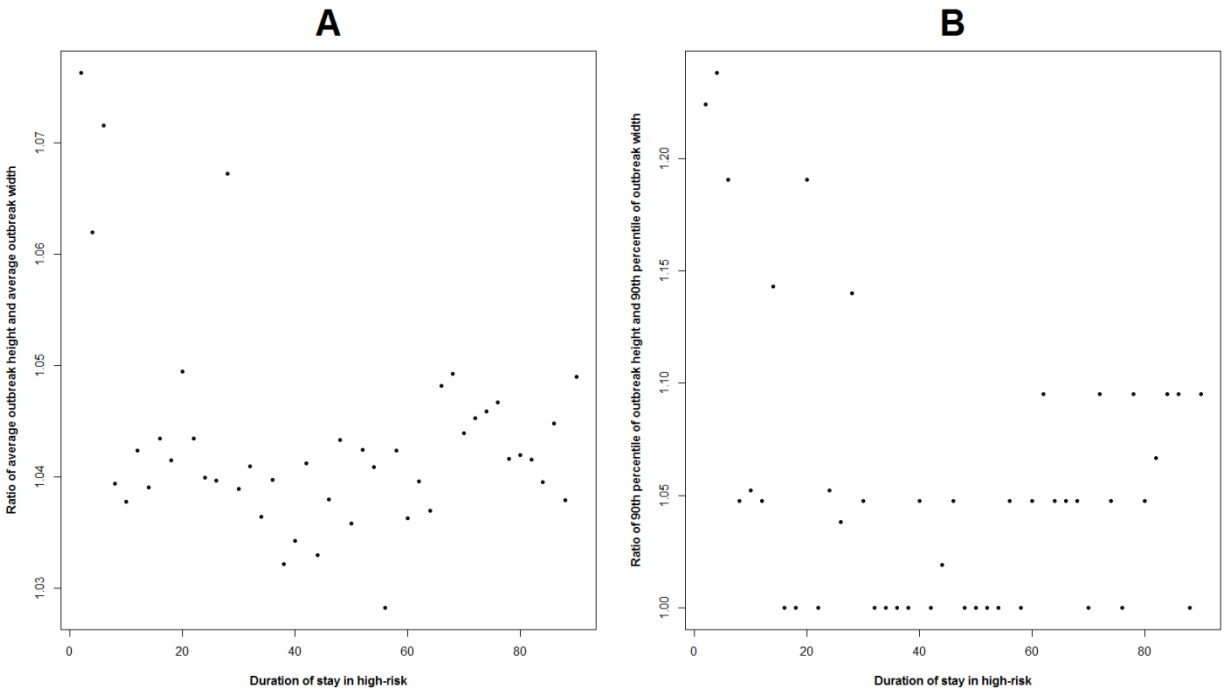

Figure C-3: 2D scatter plots with the duration of stay in high-risk on the X-axis for the data stratified with respect to 'high-to-low contact ratio': 1 and 'AHI transmission potential': 0.5. Panel A: Ratio of average outbreak height and average outbreak width. Panel B: Ratio of the 90th percentile of outbreak height and the 90th percentile of outbreak width.

Boxplots of different AHI outbreak measures with respect to the duration of stay in high-risk phase.

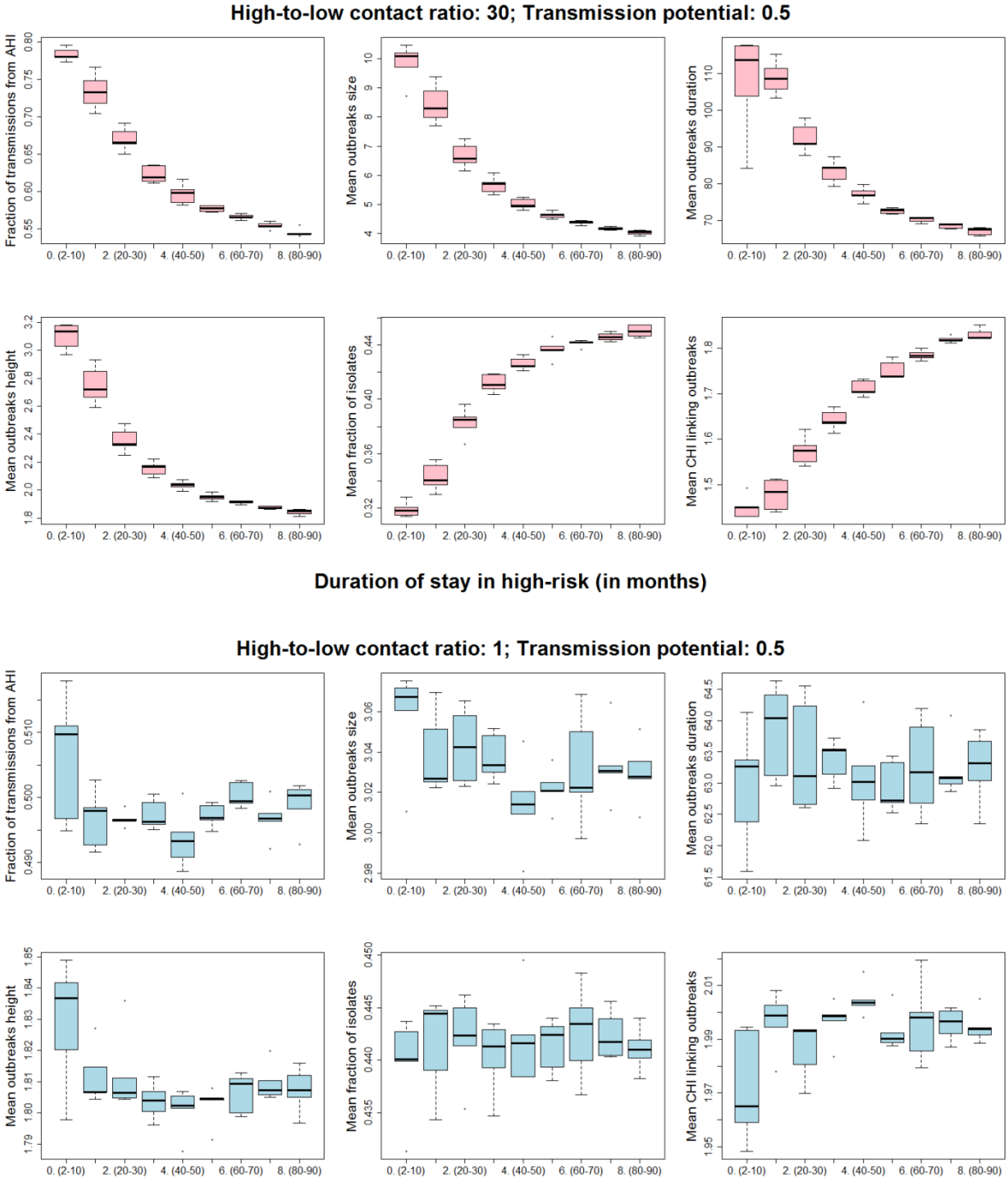

Figure C-3: Box plots for AHI outbreak measures with respect to the 'duration of stay in high-risk' parameter. Data stratified with respect to the 'high-to-low contact rate ratio' (rCHL). From top left in row-major order: fraction of transmission from AHI, mean outbreak size, mean outbreak duration, mean outbreak height, mean fraction of isolates and mean chronic links (CHI) connecting AHI outbreaks. AHI Transmission Potential: 0.5. FrH: 0.1. FHatH: 0.9. Medians are denoted by solid black lines while the top and bottom box edges denote the first and third quartile. Whiskers denote the largest and smallest data within 1.5 times the interquartile range. Above: high-to-low contact ratio: 30. Below: high-to-low contact ratio: 1.

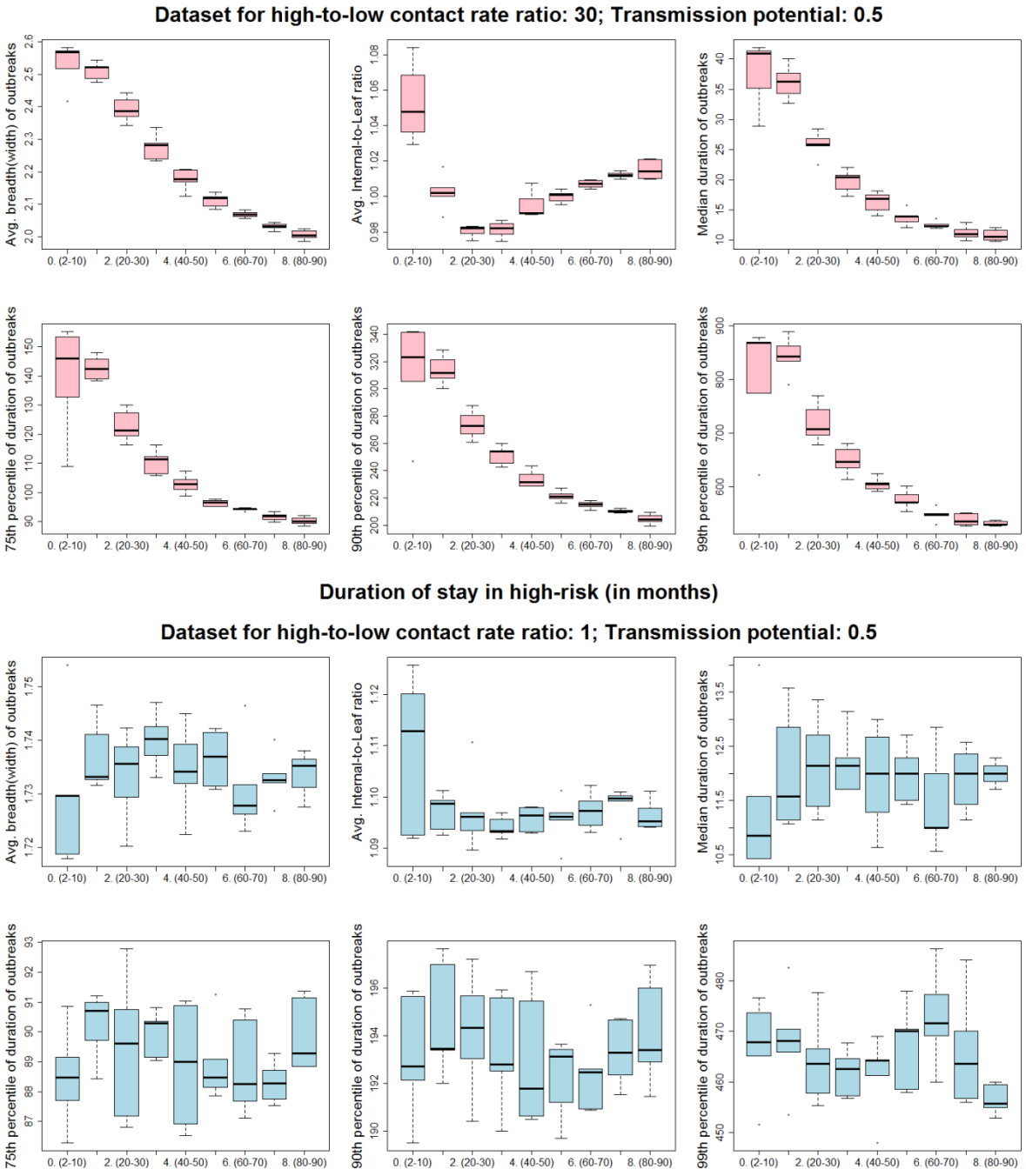

Figure C-4: Box plots for AHI outbreak measures with respect to the 'duration of stay in high-risk' parameter. Data stratified with respect to the 'high-to-low contact rate ratio' (rCHL). From top left in row-major order: mean width (maximum secondary transmission by an individual in an AHI outbreak), mean internal-to-leaf nodes ratio, median duration of outbreaks in days; and 75th, 90th and 99th percentiles of duration in days, respectively. AHI Transmission Potential: 0.5. FrH: 0.1. FHatH: 0.9. Medians are denoted by solid black lines while the top and bottom box edges denote the first and third quartile. Whiskers denote the largest and smallest data within 1.5 times the interquartile range. Above: high-to-low contact ratio: 30. Below: high-to-low contact ratio: 1.