

# Chapter 22

## Control of Epidemics on Hospital Networks

Vitaly Belik, Philipp Hövel and Rafael Mikolajczyk

**Abstract** The spread of hospital-related infections such as antibiotic-resistant pathogens forms a major challenge in public healthcare systems world-wide. One of the driving mechanisms of the pathogen spread are referrals or transfers of patients (hosts) between hospitals or readmissions after their stay in the community, constituting a dynamical network of hospitals. We analyze referral patterns of 1 million patients from one Federal State in Germany over the period of three years. We extract the underlying statistics of relocation patterns and build an agent-based computational model of pathogen spread. We simulate an outbreak of an SIS-type infection (susceptible-infected-susceptible) and evaluate characteristic time scales and prevalence levels. For such recurrent diseases, we finally investigate the effect of control measures based on screening and isolation of incoming patients.

### 22.1 Introduction

The emergence and transmission of antibiotic-resistant pathogens is an issue of a major challenge for public health on a world-wide scale [1]. Due to the availability of data and computational resources, a number of investigations have been devoted to study the pathogen spread in hospital networks in different countries [2, 3]. It turned out that countries differ in their hospital network structure [2]. Therefore, it is important to analyze healthcare systems in different countries to understand

---

V. Belik (✉) · P. Hövel  
Institut für Theoretische Physik, Technische Universität Berlin,  
Hardenbergstraße 36, 10623 Berlin, Germany  
e-mail: belik@mailbox.tu-berlin.de

V. Belik · R. Mikolajczyk  
Helmholtz Centre for Infection Research, Mascheroder Weg 1,  
38124 Braunschweig, Germany  
e-mail: Rafael.Mikolajczyk@helmholtz-hzi.de

P. Hövel  
Bernstein Center For Computational Neuroscience Berlin,  
Philippstraße 13, 10115 Berlin, Germany  
e-mail: phoevel@physik.tu-berlin.de

universal features and heterogeneities. Studies that analyze the German healthcare system from the network perspective are very rare [4]. We aim to fill this gap and present an analysis of German data and for the first time model the spread of a pathogen on this network of hospitals in Germany. Specifically, we elaborate on the impact of screening procedures of patients admitted to hospital to reduce the prevalence level of a disease. Additionally we incorporate the possibility of patients carrying the pathogen after their release to the community, which was not considered in the previous studies [2, 3]. Our study should be considered as a proof of concept for approaches combining complex network theory and computational methods in epidemiology.

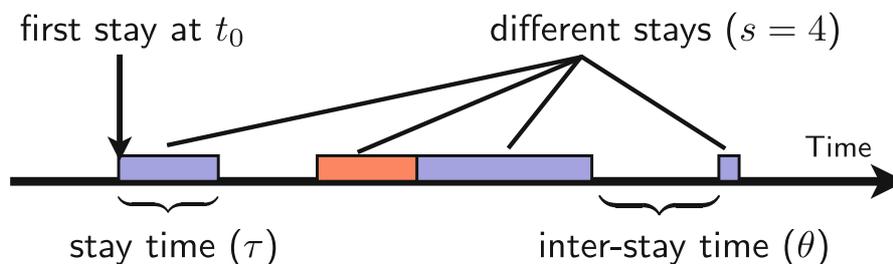
The rest of this chapter is organized as follows: In Sect. 22.2, we introduce the dataset and provide details of the model. We also present an analysis of the dataset in terms of its network and temporal properties. Section 22.3 contains the main numerical results and discuss the influence of screening procedures of patients upon admittance to a hospital. We finally summarize our findings in Sect. 22.4.

## 22.2 Dataset and Model

In the presented study, we consider anonymized data on patient referrals, that is, relocations between hospitals or release to/readmission from the community. The dataset contains 1654 hospitals, which are considered as nodes in the network,  $9.18 \cdot 10^5$  patients with around 2 million hospital stays over the course of 3 years of data.

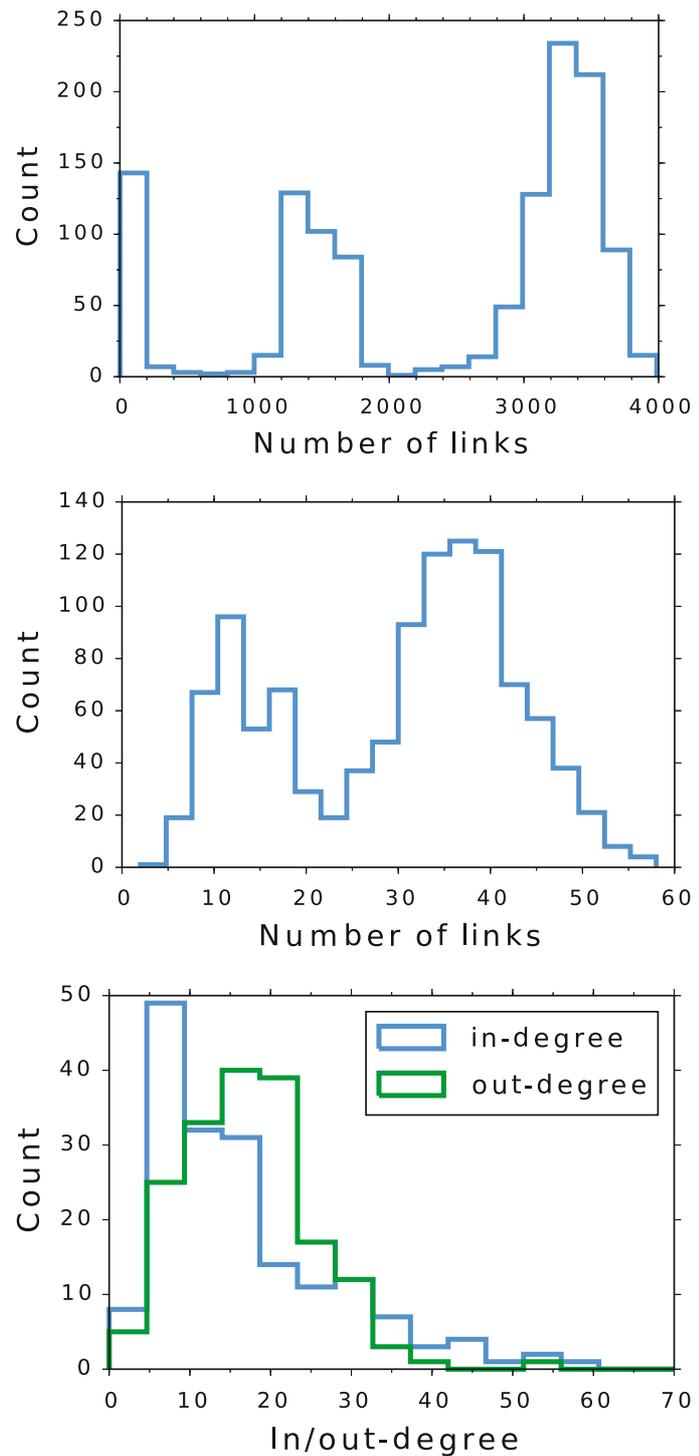
The data was obtained from a healthcare provider in a large federal state in Germany. It contains the following information about the referral: day of first admission  $t_0$ , number of stays  $s$ , duration of each stay  $\tau$ , and inter-stay time  $\theta$ . See Fig. 22.1. The color corresponds to different hospitals. Patients can be directly transferred between hospitals or spend some time in the community.

On an average week day there are around 3400 relocation events as shown in Fig. 22.2 (top). These relocations form the set of links in our network. Note that the sequence of links is crucial to ensure the causality of a spreading process. In

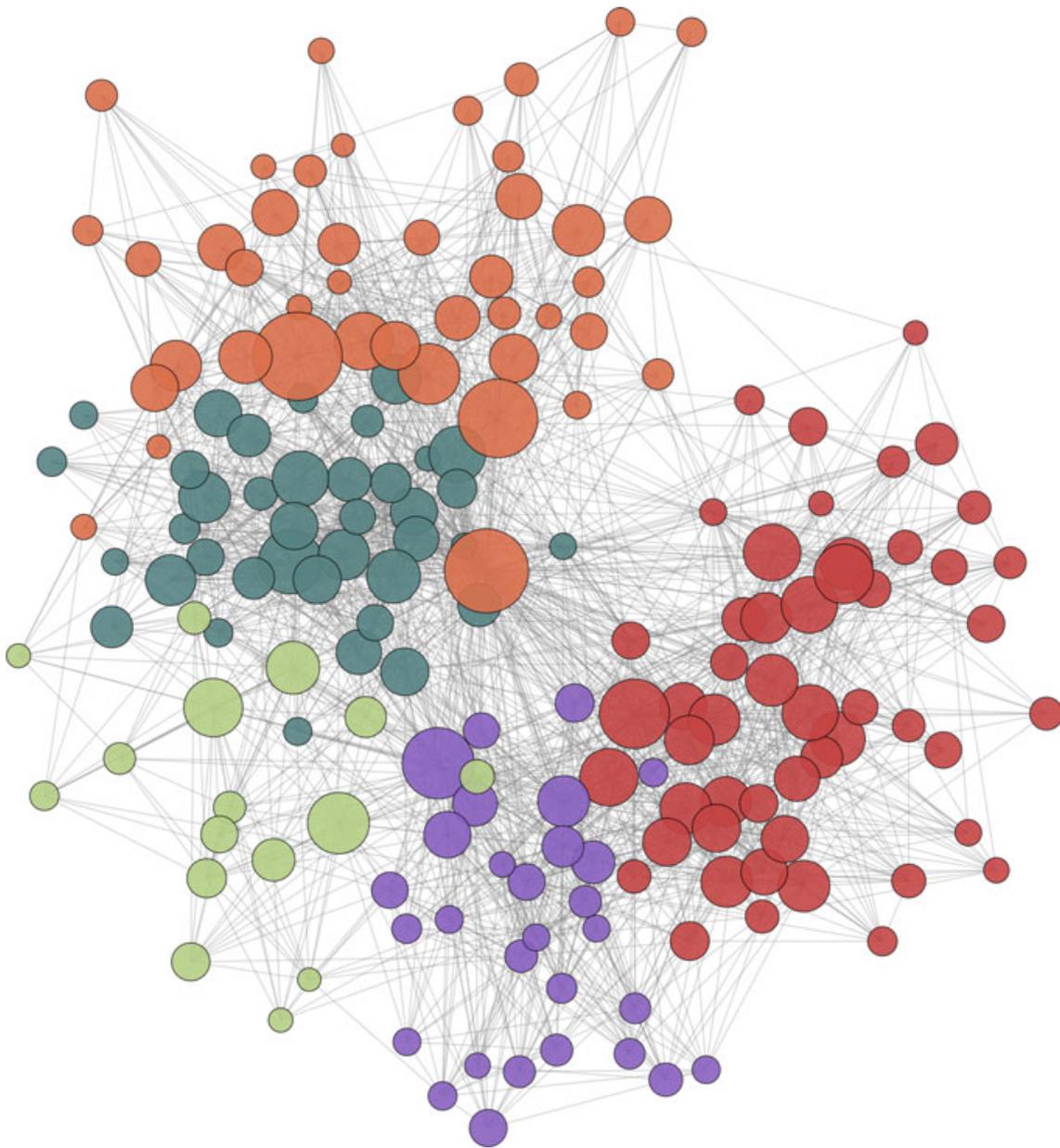


**Fig. 22.1** Schematic of the available referral data for an exemplary patient with the day of first admission  $t_0$ , number of stays  $s$ , duration of stay  $\tau$ , and inter-stay time  $\theta$ . The *color* indicates different hospitals

**Fig. 22.2** Properties of the time-aggregated, undirected, and non-weighted hospital network with major 176 nodes depicted in Fig. 22.3. *Top* Histogram of the number of all relocations per day. *Middle* Direct relocation between pairs of hospitals. *Bottom* In- and out-degree histograms



other word, the network under investigation constitutes a temporal network [5, 6]. If we consider only direct relocations between pairs of hospitals without relocations between hospitals and homes (community), we observe around 40 relocations on an average week day. See Fig. 22.2 (middle). The in- and out-degree distributions of the aggregated network of hospitals are presented in Fig. 22.2 (bottom). The in-degree is broader distributed than the out-degree. Note that there is one outlier referring to a node with in-degree 120 (not shown).



**Fig. 22.3** Modular structure of the patient transfer network. 176 nodes (hospitals) can be subdivided into 5 modules indicated by the *color* of the node. The position is chosen according to a spring-embedded layout and thus, corresponds to the topological position in the network and not to the geographical location. Node sizes corresponds to hospital sizes (number of beds) as estimated from the data

The data might also contain information about referrals to hospital outside the federal state under consideration, but does not include full referral records beyond the state borders. We identified the hospitals located in the state under consideration as those hospitals with a maximum number of patients larger than a threshold which was set to 30. This resulted in 176 frequently visited hospitals as depicted in Fig. 22.3. In this reduced network we computed 4 modules using the Louvain algorithm described in Ref. [7]. This method sorts all nodes into different modules by maximizing the so-called modularity  $Q$ , which is defined as

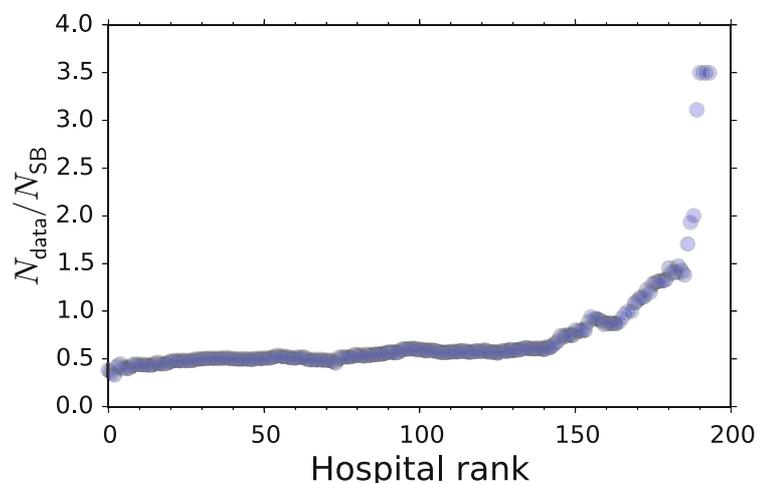
$$Q = \frac{1}{2m} \sum_{i,j=1}^N \left( a_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j), \quad (22.1)$$

where  $m$  denotes the total number of links,  $\{a_{ij}\}$  is the adjacency matrix, and  $k_i$  and  $c_i$  refer to the degree and module of node  $i$ , respectively. To obtain Fig. 22.3, we applied this algorithm to the time-aggregated, undirected, and non-weighted network. Information on the modular structure of networks can be used to identify critical links for an effective prevention of epidemics. This is important, for instance, to contain an outbreak locally and prevent spreading across different modules [8].

Due to reasons of privacy we do not have information on geographical coordinates of hospital or access to other types of metadata. However, we are able to estimate hospital sizes from the data. For this purpose we take the maximum number of patients sojourning in the individual hospital as an estimate of its size, i.e. number of beds. To verify our estimates, we compare these with the data from statistical bureau by ranking both numbers.<sup>1</sup> The accuracy of this procedure is shown in Fig. 22.4. We find that the ratio of the ranked hospital size, which are estimated from the data, and the real hospital sizes remains constant around 0.5 for the first 150 nodes. Assuming a uniform distribution of the customers of the health insurance company, which provided us with data, this ratio nicely reflects its market share as it is close to the real value of around 40%. The agreement can be further improved, if we take into account an occupancy rate of hospitals below 100%.

Quantifying categories of links in our network, we also computed the numbers of relocations between community and hospitals ( $2.974 \cdot 10^6$  or 99% of all relocations) and direct transfers between different hospitals ( $3.3 \cdot 10^4$  or 1% of all relocations). Therefore, it is to be expected that the role of the community is very important, as the majority of patients are not directly transferred between hospitals, but first stay for some time in the community, potentially carrying the pathogen. In the simulations presented below, we assume no disease spreading in the community for simplicity,

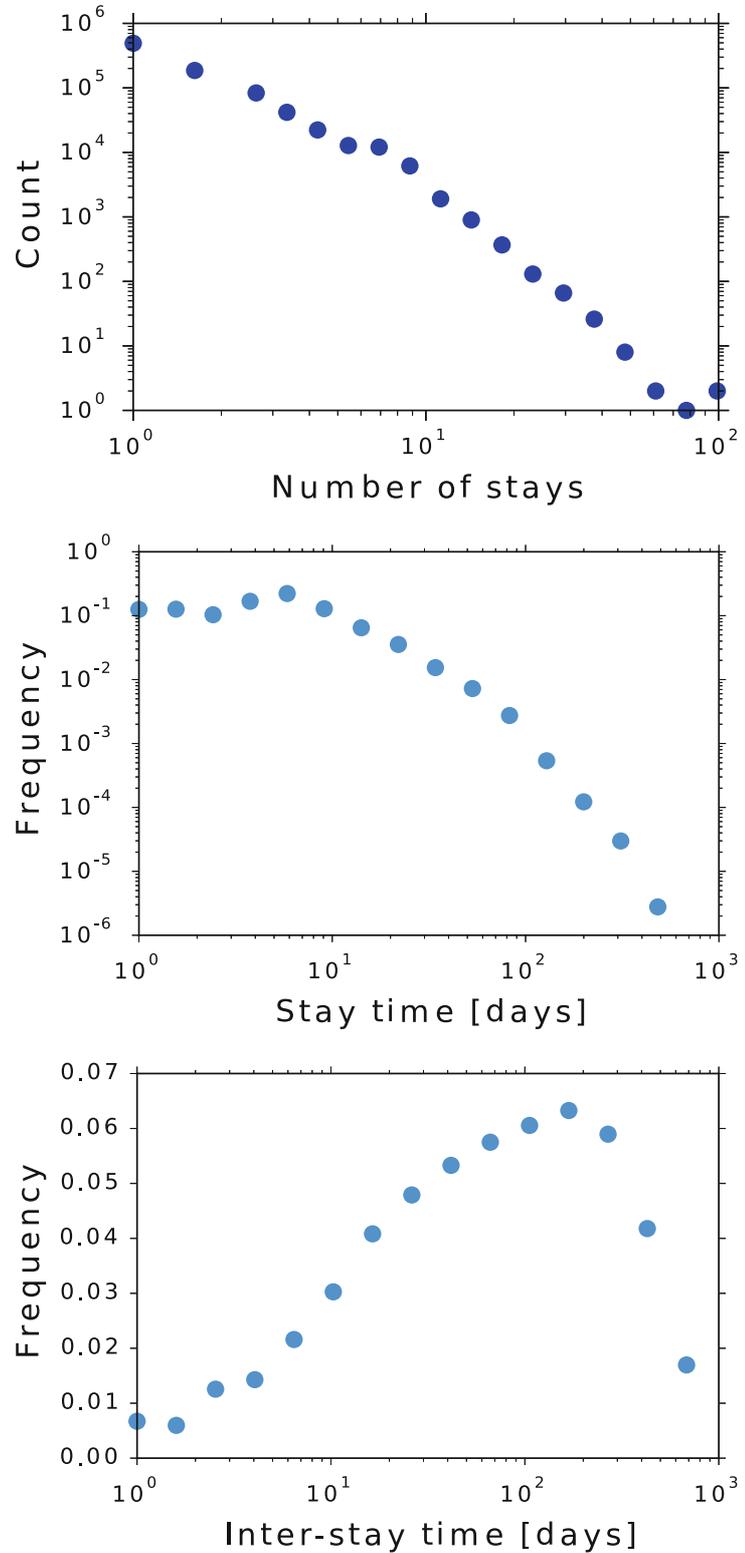
**Fig. 22.4** Evaluation of the robustness of hospital size estimations from the data.  $N_{SB}$  are the number of beds in hospitals as given from the statistical bureau.  $N_{data}$  are the sizes of hospitals as estimated from the data, i.e. the maximal number of patients on any day



<sup>1</sup>See *Landesamt für Statistik Niedersachsen*: <http://www.statistik.niedersachsen.de>.

because we lack information about the moving patterns except for the times of hospital (re)admission and release. Since we do not have any information about the internal ward structure of the hospitals either, we assume that within a hospital patients are well mixed, and the law of mass action holds.

**Fig. 22.5** Statistics of the dataset: histograms of (*top*) number of stays, distribution of (*middle*) duration of stays, (*bottom*) distribution of inter-stay times



**Fig. 22.6** Total number of agents in the system in dependence on time

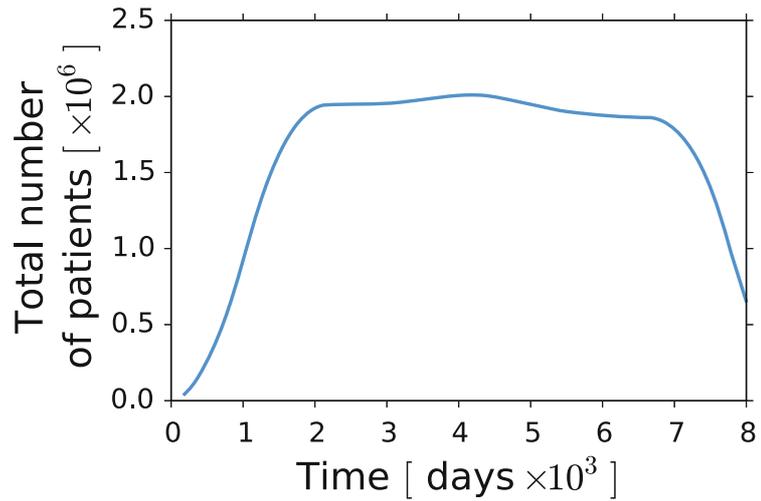


Figure 22.5 depicts basic statistics of the individual referrals in the considered dataset. See also schematics shown in Fig. 22.1. The number of stays and stay times (top and middle panels in Fig. 22.5) are broadly distributed with long tails. The majority of the patients (around 50 %) were admitted just one time to a hospital. The inter-stay time distribution peaks around 110 days (bottom panel in Fig. 22.5).

In order to simulate the whole population of size  $N$ , we generate the corresponding number of agents from the data on only 40 % of the patients using the following procedure: (i) we chose  $N$  times a patient ID from the dataset with its referral record and assign a random day of its first appearance in the interval  $[0, T]$  with  $T = 3$  a, (ii) we periodically repeat these records for the intervals  $[(n - 1)T, nT]$  with  $n = 1, 2, 3$  leading to a total observation time of 12 a; (iii) taking into account mortality rates for the agent, an agent is removed assuming a death rate of  $0.007/a$  and at the same time, a new agent with the same referral profile is added. Following this procedure, we reach a constant population level after an initial transient as shown in Fig. 22.6.

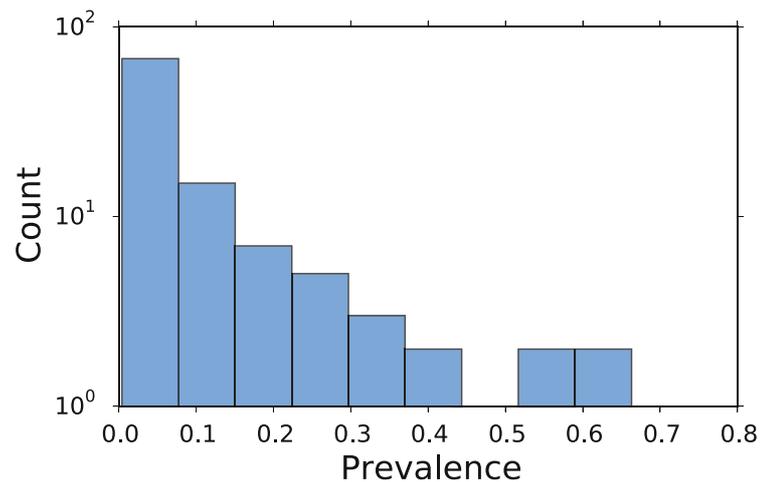
In order to model an endemic prevalence level of resistant pathogens, we consider an SIS (susceptible-infected-susceptible) epidemic model. Given the number  $I$  of infected and  $S$  of susceptible individuals in one hospital, the dynamics follow a chemical kinetic equations for infectious dynamics:



where  $\alpha$  and  $\beta$  denote the infection and recovery rate per individual. Thus we consider the frequency-dependent model, where the chance of infection is proportional to the product of the number of susceptible and infected individuals in a single population and inversely proportional to its size. Equation (22.2) describes an undetected, free-running spread of pathogens in the absence of control measures.

We use a stochastic agent-based computational epidemic model on a network of hospitals and implement the events according to the empirical data using a priority queue data structure [9] to keep track of single individuals, their infection status and

**Fig. 22.7** Histogram of the time-averaged endemic prevalence values in hospitals with a non-zero number of infected agents. Infectious rate  $\alpha = 0.1/\text{day}$  and recovery rate  $\beta = 2.7 \cdot 10^{-3}/\text{day} = 1/\text{year}$



time events (arrival at and release from a hospital, recovery). For the time between two subsequent events, the local node dynamics follow the SIS-model described by the kinetics (22.2). The infectious rate  $\alpha$  is chosen to ensure the average prevalence within hospitals around 5%. We consider a recovery rate  $\beta = 1/a = 2.7 \cdot 10^{-3}/\text{d}$ , which corresponds to typical carriage times of a bacterial pathogen. We configure the system in the following way: we populate the hospitals according to the procedure described above using the empirical transfer profiles of the dataset. As initial condition, we implement 0.99% of the patients in all hospitals as infected. We find that the dynamics of our network model reach an endemic state after 1000 days (Fig. 22.4).

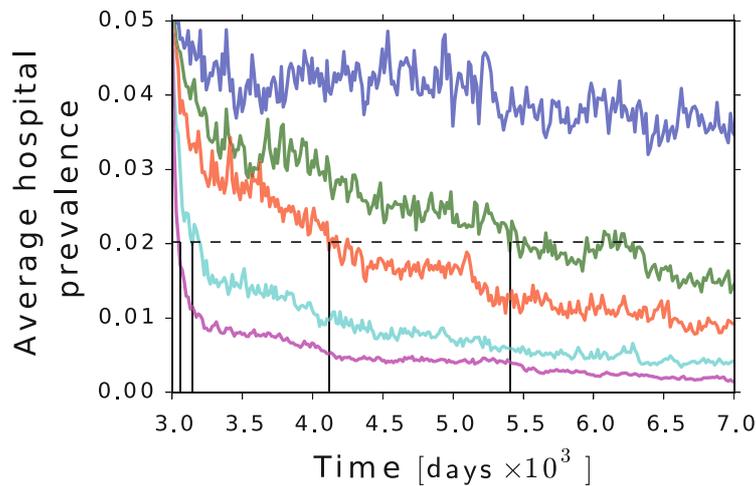
Figure 22.7 depicts the histogram of the time-averaged endemic node prevalence. The median is 0.02 and the mean value is 0.07. One can see that the prevalence distribution is inhomogeneous and skewed towards low values, indicating a small prevalence for many nodes.

## 22.3 Results of Simulations and Control by Screening

Extending the model described in the previous section, we additionally implement the following control measure. We randomly screen a fraction  $\nu$  of patients incoming in every hospital. Assuming a test sensitivity of 100%, patients that are detected as infected are immediately isolated and cured from the disease.

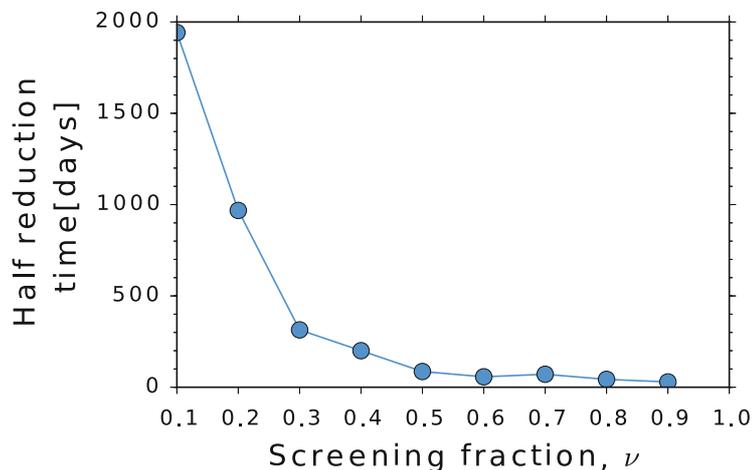
Figure 22.8 presents the time series of the prevalence after the control is applied at  $t = 3000$  d for different screening fractions  $\nu$ . Note that  $\nu = 0$  corresponds to the uncontrolled case. As intuitively expected, screening leads to a reduction of the prevalence level. We observe that the screening rate has to be considerably high in order to achieve significant results. For a 10-fold reduction within 300 days, for instance, a screening of 90% of incoming patients is required.

Figure 22.9 shows the time required to reduce the prevalence to 50%, which is known as *half reduction time*, in dependence on the screening fraction  $\nu$ . This half reduction time is marked in Fig. 22.8 by vertical lines. We find that the half



**Fig. 22.8** Averaged prevalence for different fraction of incoming patients screened:  $\nu = 0$ , i.e. no screening, (*blue*),  $\nu = 0.1$  (*green*),  $\nu = 0.2$  (*red*),  $\nu = 0.5$  (*cyan*),  $\nu = 0.9$  (*magenta*). The average is computed over all hospitals for each time step. The *dashed horizontal line* indicates the 50% reduction level and the vertical lines mark the half reduction times. System parameters of the SIS model (22.2) as in Fig. 22.7

**Fig. 22.9** Time until 50% prevalence reduction from the start of the screening. System parameters of the SIS model (22.2): infectious rate  $\alpha = 0.1/\text{day}$  and recovery rate  $\beta = 1 \text{ year}^{-1}$



reduction time decreases strongly with screening rates up to  $\nu = 30\text{--}40\%$ . For higher values of  $\nu$  the half reduction time equilibrates around 100 days and does not change significantly. Thus, if the goal is to reduce prevalence to 50%, moderate screening fractions  $\nu = 30\text{--}40\%$  are sufficient.

## 22.4 Conclusion

We have shown how complex network theory and computational methods of agent-based stochastic reaction-diffusion processes can help to control nosocomial infectious diseases, such as antibiotic-resistant pathogens, e.g. Methicillin-resistant *Staphylococcus Aureus* (MRSA) or *Clostridium Difficile*. We have analyzed patients referral patterns in one federal state in Germany over the period of 3 years. We have extracted the corresponding hospital network of patient relocations and built a com-

putational agent-based stochastic model of disease dynamics including the full history of hospital stays on the single patient level. We have assessed the efficiency of screening a fraction of incoming patients as a potential control measure. This means that in the case of a positive screening test, the patient is isolated and cured before admission to the hospital. For typical values of parameters, we have found that the endemic prevalence can be halved within 100 days for screening fraction around 30–40 %.

Our study represents a proof of concept and opens roads for the future analysis of the potential impact of different epidemic control measures in a network of hospitals.

**Acknowledgments** All the authors acknowledge the courtesy of the AOK Niedersachsen for providing the anonymized data on patient referrals. VB thanks André Karch and Johannes Horn for fruitful discussions. VB and PH acknowledge funding by the Deutsche Forschungsgemeinschaft in the framework of Collaborative Research Center 910. At the early stage of this study VB was financially supported by the fellowship “Computational sciences” of the VolkswagenStiftung.

## References

1. World Health Organization, *Antimicrobial Resistance: Global Report on Surveillance*. WHO, Geneva (2014)
2. T. Donker, J. Wallinga, R. Slack, H. Grundmann, PLoS ONE **7**(4), 1 (2012). doi:[10.1371/journal.pone.0035002](https://doi.org/10.1371/journal.pone.0035002)
3. J.F. Gracia, J.P. Onnela, M.L. Barnett, V.M. Eguíluz, N. Christakis, arXiv preprint [arXiv:1504.08343](https://arxiv.org/abs/1504.08343) (2015)
4. M. Ciccolini et al., Int. J. Med. Microbiol. **303**(6–7), 380 (2013). doi:[10.1016/j.ijmm.2013.02.003](https://doi.org/10.1016/j.ijmm.2013.02.003)
5. P. Holme, J. Saramäki, Phys. Rep. **519**, 97 (2012). doi:[10.1016/j.physrep.2012.03.001](https://doi.org/10.1016/j.physrep.2012.03.001)
6. A. Casteigts, P. Flocchini, W. Quattrociocchi, N. Santoro, Int. J. Parallel Emergent Distrib. Syst. **27**(5), 387 (2012). doi:[10.1080/17445760.2012.668546](https://doi.org/10.1080/17445760.2012.668546)
7. V.D. Blondel, J.L. Guillaume, R. Lambiotte, E. Lefebvre, J. Stat. Mech. **10**, P10008 (2008)
8. N. Masuda, New J. Phys. **11**(12), 123018 (2009)
9. T.H. Cormen, *Introduction to Algorithms*. MIT Press (2009)