

Model Alignment of Anthrax Attack Simulations¹

Li-Chiou Chen¹, Kathleen M. Carley¹, Douglas Fridsma², Boris Kaminsky¹, Alex Yahja¹

¹ Institute for Software Research, International

School of Computer Science

Carnegie Mellon University

² Center for Biomedical Informatics

School of Medicine

University of Pittsburgh

Abstract

This paper describes our experience aligning two simulation models of disease progression after biological attacks. The first model is the Incubation-Prodromal-Fulminant (IPF) model, a variation of the Susceptible-Infected-Recovered (SIR) epidemiological model, and the second is an agent-based model called BioWar. We run BioWar simulations to see whether the results will, at the population level, match the IPF results. We showed that BioWar can generate population level results that are close to IPF. In addition, BioWar outputs emergent properties that cannot be simulated in IPF. This study provides insights for modelers who are developing simulation tools for investigating bioterrorism attacks and for decision makers who use these tools.

Keywords: model alignment, anthrax, bioinformatics, agent-based model, biosurveillance, simulations.

¹ Corresponding author: Li-Chiou Chen, 231 Smith Hall, Institute for Software Research, International, Carnegie Mellon University, Pittsburgh, PA 15213. Email: lichiou@andrew.cmu.edu. TEL: 412-2687527.

Model Alignment of Anthrax Attack Simulations

1. INTRODUCTION

To make informed decisions on how to respond to bioterrorism, policy analysts need to include the complex social responses and disease processes inherent in bioterrorism attacks. We are developing an agent-based simulation model (BioWar) to aid the decision making process. BioWar is a simulation tool that combines computational models of social networks, communication media, disease models, demographically resolved agent models, spatial models, wind dispersion models, and a diagnostic model into a single integrated system that can simulate the impact of a bioterrorist attack on a city [10, 23]. In BioWar analysts can model real cities using census, school district demographics, and other publicly available information.

Disease processes and response strategies are traditionally modeled by the susceptible-infected-recovered (SIR) model. The SIR model and its variations have been widely used to model the spread of epidemics and to study immunization strategies [1, 3, 12]. The SIR model is a “population-based” description of disease progression processes that assume homogeneous mixing of individuals. The agent-based BioWar takes a different approach thus allowing us to model the complex social interactions absent in most SIR models. However, in order to understand the benefits and limitations of using BioWar to model biological attacks, we aligned BioWar with a population-based model revised from the SIR model. This process is called model alignment.

Model alignment [2], also referred to as “docking,” is the comparison of two computational models to see if they can produce equivalent results. Properly done, model

alignment can uncover the differences and similarities between models and reveal the relationships between the different models' parameters, structures, and assumptions. By aligning a complex new model with a simpler and well-understood model, one can obtain a sense of validity needed to develop the new model. The same technique has been used previously to validate a model of organization performance [18]. This study is a part of a greater validation process for BioWar [10, 11]. Our purpose is to demonstrate a general equivalence between BioWar and SIR based on anthrax attack simulations.

To calibrate the revised SIR model and some BioWar parameters, we used empirical data sets based on known release of aerosolized anthrax spores. Since anthrax is not contagious, we have to revise the original SIR model. We used the revised model as an instrument to examine the predictions from BioWar and to investigate the factors causing the differences and similarities between the predictions.

This paper is organized as follows. Section 2 provides background information on BioWar and the revised SIR model, and compares these two models qualitatively. Section 3 explains the processes of model alignment. Section 4 compares BioWar and the revised SIR model based on simulation results on the release of aerosolized anthrax spores. In addition, this section discusses what can be improved in BioWar based on the results. Finally, conclusions on the contributions and future works are in section 5.

2. THE TWO MODELS

BioWar models the residents of a city (agents) as they go about their lives. When a bioattack occurs, those in the vicinity of the release may become infected, following probabilistic rules based on received dose and age of the agent. The infected agents modify their behaviors as their disease progresses and they become unable to perform

their normal functions as the disease worsens. A detailed description of the model along with a plan for validation and preliminary validation results can be found in [10]. In this paper only the anthrax attack and disease progression simulation is discussed.

In principle, agent-based models have the advantage that the heterogeneity of individual response can be accounted for, thus enabling a finer grained analysis and allowing the tools to be used for training and intelligence purposes. In BioWar, a further advantage is that the diseases are modeled at the symptom level thus enabling the model to contribute to our understanding of the ways in which early symptomatic based behaviors, such as the purchase of the over-the-counter-drugs are likely to emerge after a biological attack. Further, by using a general symptom based framework, new diseases and even “unheard of” diseases can be rapidly modeled in BioWar. Additionally, in BioWar, multiple diseases are simultaneously tracked so that disease interactions can be examined.

In contrast, the susceptible-infected-recovered (SIR) model assumes a homogeneous population and is typically instantiated for only a single disease at a time in terms of response states rather than symptoms. Nevertheless, the SIR model has been a widely adopted model of the spread of a disease through a population. As noted, the SIR model is a population-based description of the epidemic diffusion process that categorizes the entire population into three states: susceptible (S), infected (I) and recovered (R). The SIR model assumes that the population is homogeneous. That is, all members of a particular state are identical and have predefined transition probabilities of moving to another state in the model. Although variations in the way in which the disease is manifested and symptom based behaviors can be tracked using Monte Carlo

simulation methods, the interaction among population members is often lost. Further, in an SIR model, modeling the impact of a multiple diseases on a population creates unmanageable complexity in the models and limits the value of any one model for the study of multi-disease attacks. Most SIR models are not spatial models, only recently does work on spatial-epidemiology progress [19].

These comments aside, there are some critical advantages to SIR models. First, they are widely used and understood by the medical and policy-making community. Secondly, once the transition probabilities for a disease are known, an SIR model can be rapidly developed. Third, SIR models are relatively easy to link to economic cost models thus enabling first order cost-benefit analyses to be conducted.

Looking at specific examples one can see additional similarities and differences of these models. Using anthrax attacks as an example, we compare BioWar with a population-based model that is derived from the SIR model. We chose anthrax attacks as an example because of the need to study response strategies against large scale weaponized attacks, of which anthrax is one of the most likely candidates.

It should be noted that inhaled anthrax is infectious but is not contagious, so we revised the SIR model. We call the revised SIR model an IPF model (Figure 1), because it distinguishes between the three stages of anthrax disease progression: incubation, prodromal, and fulminant. Similar models have been used to estimate medical costs of anthrax response systems [7]. The revised model is a Markov model, in which state variables (represented as rectangular boxes in Figure 1) are populations in a certain disease stage and transition probabilities determine the population flow (represented as arrows) from one state to another. Appendix A describes the model mathematically.

At the beginning of an attack, we simulate the release of anthrax spores over a city on a specific day, exposing some fraction (E) of the population. A fraction of these will become infected after inhaling anthrax spores and start the three stages of the disease progression. Incubation (I) refers to the fraction of the population that is infected by anthrax spores but has not shown any symptom yet. Prodromal (P) refers to the fraction of the population that shows a spectrum of non-specific symptoms such as fever, chills, cough and vomiting. Fulminant (F) refers to a fraction of the population who develops symptoms abruptly, with sudden fever, dyspnea, diaphoresis and shock, or more specific and severe symptoms [5, 15]. For each of the three states, some persons may be treated and enter either one of the other three treatment states (ITX, PTX and FTX) representing treatment in hospitals. For each of the six disease states, people have a certain probability either recovering (recovery state, R) or dying (death state, D).

Qualitatively speaking, the differences between IPF and BioWar can be summarized as follows:

- *Population assumptions:* IPF models population cohorts as they transition through different disease states, i.e., same number of social contacts. BioWar models heterogeneous individuals and their interaction in social settings. i.e., various numbers of social contacts as agents go about their daily life. Moreover, BioWar individuals have spatial locations. For example, children go to schools that are in the districts near their homes.
- *Disease model design:* IPF simulates the disease progression from a macro point of view. That is, the model uses a state machine to describe the state changes among sub-populations and uses proportional state transition

probabilities to describe the migration of sub-populations. BioWar simulates the emergent properties of individual agents from a micro point of view. That is, to describe the population level disease status, BioWar models and summarizes the disease state of the individual agents. The macro behavior of the population emerges from the outcomes for the individual agents. For example, IPF models the population in incubation stage having a transition probability to move to the symptomatic stage while BioWar models each agent having an incubation stage duration.

- *Computational process:* To generate the prevalence of a disease over time, the BioWar model requires more computational power than does IPF. In addition to tracking the maliciously introduced infection in exposed agents, BioWar models behaviors and information used in early detection algorithms as well as health status information, i.e., 60 common diseases that create the background against which bioattacks must be detected.
- *Initialization:* BioWar is initialized with information that describes individual differences. For example, agents live in different school districts and have different ages based on census data. IPF requires initial state characterization and state transition probabilities of the population. The entire population is divided into several sub-populations according to the disease stages.
- *Parameterization:* While IPF takes both the exposed population and infected population as inputs, BioWar can calculate them as emergent properties from simulating parameterized attacks. For example, BioWar can be parameterized

to describe different attack scenarios with different wind speed, release location, efficiency of the release and mass of bioagent.

3. THE PROCESS OF MODEL ALIGNMENT

We aligned BioWar with IPF and compared the outputs from both models. The results are also compared with empirical data to obtain a sense of validity for our scenarios. Figure 2 shows the process of model alignment.

First, we constructed two empirical data sets based on literature of previous anthrax releases. The first data set is based on the 2001 anthrax letter incidents in the United States [4-5, 13, 15-17]. The US data has eleven confirmed inhalational anthrax cases and five deaths even after medical treatment. The second data set is based on the 1979 anthrax outbreak in Sverdlovsk [8, 20], a part of the former Soviet Union. The Sverdlovsk data has 77 confirmed inhalational anthrax cases with 66 deaths. Appendix B describes the two data sets in detail. Based on the two data sets, we calibrated state transition probabilities of the IPF model by fitting incubation period, the number of deaths, and the number of recovered persons.

Second, we aligned the components of two models based on the stages of disease progression and developed a scenario of a large-scale anthrax attack. Finally, we compared the two models using two methods. The first order analysis, described in Section 4.1, compared the final outputs of a simulated attack, including infection rate, death rate and stabilization time (the time after which there are no new cases or deaths from the bioattacks). The second order analysis, described in Section 4.2, compared the dynamics of three sub-populations over time. Our purpose was to compare the

predictions of the two models through the first order analysis and to examine the longitudinal dynamics in these two models through the second order analysis.

3.1. Alignment of model components

In order to compare BioWar and IPF based on an identical set of inputs, we first tuned the model parameters in both models to be as close as possible. Since IPF is structurally different from BioWar, they do not share the same model parameters. Table 1 compares the differences in structure between the two models based on the stages of disease progression. For each infected agent, BioWar has a disease stage corresponding to one in IPF.

IPF takes exposed population as an input parameter and calculates the number of infected once after an attack based on the two empirical data sets. The attack model in BioWar takes input parameters such as wind speed, release height, and release mass of a biomaterial, and calculates the number of exposed and infected persons after the release of a biomaterial based on the geographical distribution of the population from census data.

Focusing only on the disease progression process of anthrax infections after people were exposed to anthrax spores, we calibrated the state transition probabilities of IPF based on the two empirical data sets. In BioWar, the disease model calculates the symptom progression of infected agents based on assumptions from disease studies and the decision model simulates the behavior of agents seeking for medical care based on medical data. The decision model decides if an agent will die or recover based on the severity of symptoms and takes into account the death rate for the disease.

3.2. BioWar scenario

For this paper BioWar was configured to represent the town of Hampton Roads, Virginia. BioWar requires considerable spatial and temporal specificity in describing an attack scenario. We chose an attack scenario in which anthrax spores were released through explosion in the air 5 meters above the municipal stadium on the 4th of July, 2003. Usually by 90 days after attack the simulation achieves a steady state, i.e., infected agents have either died or recovered.

We run BioWar scenario based on lognormal distributions for disease stage durations with the mean and standard deviation estimated from the Sverdlovsk data [8, 22]. Table 3 shows model parameters and assumptions of our scenario. The attack releases 3000 grams anthrax spores. In our simulation, efficiency means the fraction of the live microorganisms survived in the aerosol form with sizes between 1 and 5 micron after the release that may happen as explosion, or spray release. We simulate explosive release in our experiments so that the efficiency is set to 0.05 [21]. Therefore the attack effectively releases 150 grams of anthrax spores. In our attack scenario, no detection or response systems are placed at either medical centers or emergency rooms. As a result, most patients who are exposed or infected by anthrax spores do not know that they are infected and do not obtain prophylactic treatment. However, once they fall seriously ill, they receive treatment according to the severity of their symptoms.

Epidemiological studies provide different opinions on whether the anthrax stage durations are dose dependent. Although statistical analysis of the Sverdlovsk case did not reveal any stage duration dose dependency [8], other studies have reported the dose dependency at least for the incubation stage [6] and it is logical to assume that the two

other stages may also be dose dependent [9]. To determine which assumption we shall adopt, we conducted a test simulation on both assumptions. The means of the lognormal distribution in the dose-independent case and the dose dependent case are shown in Table 2a and Table 2b, respectively. The standard deviations for both cases are the same and are only shown in Table 2a. We found that the system dynamics for the dose independent assumption and the dose dependent assumption are similar but dose independent assumption is slightly closer to the empirical data. For example, Figure 3 shows that mortality (the ratio of death to infected population) based on dose independent assumption is closer to the Sverdlovsk data by 10% in the first 20 days. Because of this finding, we decided to run the BioWar scenario with dose independent assumption only.

4. RESULTS AND DISCUSSION

4.1. First order analysis – death rate, infection rate, and stabilization time

We compared the results of BioWar and IPF simulations with empirical data sets (Table 4). Death rates from BioWar scenario are close to those from IPF. In addition, both BioWar and IPF death rates are comparable to the Sverlovsk data and the US case.

In BioWar, the exposed population is an emergent property (thus a simulation output), which we calculated as the number of persons who have inhaled at least one anthrax spore. In contrast, in IPF the exposed population is an input parameter, which can be taken directly from real world cases but cannot be predicted in future attack scenarios as we did for the town of Hampton Roads. However, in a real world attack, exposed population is hard to calculate because it is difficult to examine everyone and determine whether or not he/she has inhaled an anthrax spore. For calibrating IPF, we estimate the exposed population to be the number of persons who received prophylaxis

for possible exposure to anthrax spores. In the US case, 10,300 people completed the 60-day course of anti-microbial prophylaxis and, in the Sverdlovsk case, 47,200 persons were vaccinated.

Similarly, infection rate (the ratio of number of infected to the number of exposed) is also an emergent property from BioWar simulations but an input parameter in IPF. IPF takes the infection rate from the empirical cases (empirical infection rate in Table 4), in which it is 0.1% in the US case and 0.16% in the Sverdlovsk case. The infection rate in the BioWar scenario (simulated infection rate in Table 4) is 10% because the exposed population is estimated differently and the released anthrax mass was about 150 times higher (BioWar effectively released 150 grams and Sverdlovsk release was estimated at about 1 gram [20]). Taking into account the differences, infection rates in BioWar are approximately the same order of magnitude as in IPF.

Stabilization time measures when the system converges. We define it as the number of days elapsed when at least 99% of infected population either die or recover. Stabilization time is a general indicator of the timing of public health responses. IPF converges 12 days earlier than the US case and 19 days earlier than the Sverdlovsk case. BioWar converges 2 days earlier than the US case and 27 days earlier than the Sverdlovsk case. The longer stabilization time in the Sverdlovsk case may be due to the resuspension of the spores from the grounds [20], which are not part of our simulations for this paper. In this aspect, IPF exhibits less difference between the two cases but BioWar reflects the discrepancy in the empirical cases.

4.2. Second order analysis - dynamics of populations over time

We compared the infected population in the three disease stages and the death rate over time to show the dynamics of BioWar and IPF. We report the outputs relative to the number of people who were infected. Because the Sverdlovsk data did not distinguish between prodromal and fulminant stages, we use the term “symptomatic” to describe the sum of the patients in these two stages. The results of the comparisons are in Figures 4-9. Each figure compares the results from BioWar and IPF with either one of the empirical data sets.

Figures 4-5 show the population in the incubation stage as a percentage of the infected population over time. For the US case, both BioWar and IPF cannot fit the empirical data well. The discrepancy comes from the small sample size (11 cases only) and the unknown exposure date of the last case. For the Sverdlovsk case, both BioWar and IPF fit the data well. Since victims in the US case are either mail workers or people who have direct contacts with mails that contain anthrax spores, the environment setting is different from the anthrax explosion in a town simulated in BioWar. We suspect that the different environment setting has an impact on the frequencies of the exposures and the dosage of anthrax spores, which may also result in the discrepancies in the incubation period.

The infected population differs by several orders of magnitude between the two models and the empirical data sets. Since we are comparing only the dynamics of the infection for the two models, we normalized the percentage of the infected population in symptomatic stage by its maximum value to rescale the results but preserve the original curve shapes. Figures 6-7 show the normalized fractions. Both IPF and BioWar simulate

a left-skew shape similar to the US and Sverdlovsk data and a spike in 10 days similar to the US data. Neither IPF nor BioWar captures the downward slope of the curve in the Sverdlovsk case (Figure 7), which exhibits an additional peak after the highest peak. Meselson, et al. [20] suspected that it is caused by the resuspension of the spores from the grounds. This result is consistent with the result in Section 4.1, in which the stabilization time in the Sverdlovsk is longer than BioWar experiments.

Figures 8-9 show the mortality among infected population over time. For the mortality, both IPF and BioWar fit the two empirical data sets well although IPF fits the US data slightly better than BioWar because of its curve fitting nature. The result shows BioWar can capture mortality rate over time as well as IPF.

4.3. Lessons learned from validating BioWar

We verified that BioWar can generate population level results that are close to IPF's and comparable to the two empirical data sets. In this exercise, we learned three aspects in validating BioWar:

- 1) The probability distribution of the disease stage durations

BioWar randomly generates the disease stage duration of an individual agent based on a probability distribution. We verified that the lognormal distribution of disease stage duration can be used in BioWar to model individual agent. The population level results, aggregated from individual agents, are as close to the Sverdlovsk data as the population-based IPF model.

- 2) Dose dependency of anthrax disease stage progression

Using BioWar we are able to examine how the difference in dose dependency assumption impacts the mortality over time while we can only use IPF to calibrate the

empirical data. From BioWar simulations, we found that the dose dependent assumption of anthrax stage duration generates about 10% more mortality in the first 20 days after the attack than the Sverdlovsk data but results in the same mortality rate afterwards. In contrast, we found that the dose independent assumption generates mortality over time closer to the Sverdlovsk case. There are two reasons to explain the discrepancy. First, missing data in the Sverdlovsk case may skew the mortality. Second, the age distribution is different between Sverdlovsk and the town of Hampton Roads that we are simulating.

3) The impact of policy responses

BioWar uses the mean and standard deviation of the lognormal distribution estimated from the Sverdlovsk data to simulate the disease progression model of anthrax without policy response of public medical interventions. In Sverdlovsk case the massive medical intervention started about 2 weeks after emergence of first cases which was probably too late [20]. The same set of parameters does not fit the US data well, as discussed in Section 4.2, because the policy response in the US case was different from the Sverdlovsk case. The policy responses influence early medical intervention and thus reduce the mortality rate of the attack. They increase the effectiveness of the treatment and extend the duration of the symptomatic stage. In this exercise, we learned that we have to adjust not only the effectiveness of the treatment in BioWar but also the disease stage durations because the infected agents can obtain appropriate treatment. In addition to verification, we found that BioWar should implement new functionalities to simulate the effects of the early detection and response strategies against biological attacks.

4.4. Comparisons between BioWar and IPF Models

The results from both BioWar and IPF fit the Sverdlovsk data well for the disease stage durations of anthrax. Compared to BioWar, the population-based IPF model fits the US data better since the transition probabilities used to determine state transitions are tightly linked to the observed data. Once calibrated to the observed data, the IPF model can be used to examine different attack scenarios and response strategies, and determine the cost-effectiveness of these strategies. However, the IPF model is limited in the kinds of interactions it can represent. As the states and population parameters increase, the complexity of the state transitions makes these models intractable. This limits the number of interactions that can be modeled.

BioWar fits the Sverdlovsk data well because the current implementation of BioWar does not simulate public announcement of attacks. The symptomatic curve in BioWar would be an order of magnitude off from the US data if we use the same means and deviations of disease stage durations estimated from the Sverdlovsk case. The quick public announcement in the US data may result in both a lower mortality rate and a longer symptomatic stage of the surviving agents than the Sverdlovsk case because of the early medical interventions. Since the individual mortality rate is reduced in our simulation based on the US data, the discrepancy shows that public response against anthrax have extended the mean and standard deviation of the lognormal distribution for the symptomatic stage at the population level. If we tune the lognormal distribution to experimentally generate the duration of the symptomatic stage that matches the population level data, BioWar will have the potential to predict additional scenarios with

different response policies, not possible with the IPF model. These findings reflect the challenges and promises of agent-based models.

In addition to the disease progression model, BioWar provides an attack model to calculate the exposed and infected populations given a certain mass and method of anthrax release, and population model describing the demographics of the town. In contrast, IPF focuses on modeling the disease progression of the infected population. It takes the exposed and infected populations after an attack as input parameters and needs other tools to estimate these populations in advance.

From this model alignment study, we found that it is fruitful to use the IPF model as an instrument to identify the areas in BioWar that can be improved. This exercise simplifies the model development process to create a more complex model based on a well-understood and simpler model. While the IPF model simulates the historical cases in the real world, the BioWar model is expected to predict a wider range of attack scenarios and the effects of various response strategies after improvements in various aspects of the model progressing from the validation foundation built on the IPF model. Work is underway to provide empirical-data-driven automated validation for BioWar and other large-scale multi-agent systems [24].

5. CONCLUSIONS

We provided the results of aligning two models of simulating disease progression after a biological attack. The two models are IPF, a population-based model, which is a revision of the SIR model, and BioWar, an agent-based model that we are developing. We showed that BioWar can generate population level results that are as close to the two

empirical data sets as IPF. In addition, BioWar outputs emergent properties (exposed population and infection rate) that cannot be simulated in IPF.

In simulating the disease progression process after biological attacks, the major difference between the population-based IPF model and the agent-based BioWar model is the stochastic nature of the simulations. While the stochastic nature of the IPF model is determined by population level of state transition probabilities, the stochastic nature of the BioWar model lies in the emergent properties of individual agents whose behaviors and decisions are determined stochastically. The difference in the stochastic nature comes from the different assumptions, where IPF assumes that the population is homogenous and BioWar assumes the population is heterogeneous and has spatial locations. For this reason, the empirical data needed for setting model parameters are different for the two models. IPF calibrates parameters based on population level statistics of an attack and BioWar needs individual level data such as census data and geographic distribution of the population.

We found that BioWar needs to adjust its parameters for the lognormal distribution of disease stage durations and the individual mortality rate once an agent is infected in order to simulate the two different public medical interventions in the Sverdlovsk case and in the US mail attack case. We can thus use the two sets of parameters to simulate other cities to realize the effect of the two different public interventions on mortality and disease progression after an anthrax attack.

By aligning the more complex BioWar with the simpler IPF model, we located several ways to tune the parameters in the disease model in BioWar. We found this exercise helpful for developing a complex system since it helps us to pinpoint the areas

that need improving. In the future, we will continue to enhance and validate the BioWar model and apply it to other cases of biological attacks in hope of using it to develop sound response strategies against biological attacks. We note that the comparisons of results from an agent-based model with an SIR model that is calibrated to real-world data is a valuable strategy for validating the agent-based model, which, once validated can be used to make predictions at levels impossible for SIR models to address.

ACKNOWLEDGEMENTS:

The authors would like to thank Neal Altman, Dr. Elizabeth Casman, and Demian Nave for their support on this paper.

This research was supported, in part, by DARPA for work on Scalable Biosurveillance Systems, the NSF IGERT9972762 in CASOS, the MacArthur Foundation, and by the Carnegie Mellon Center on Computational Analysis of Social and Organizational Systems. The computations were performed on the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputing Center. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA, the National Science Foundation, the Pittsburgh Supercomputing Center, the MacArthur Foundation, or the US Government.

REFERENCES

- [1] R. M. Anderson and R. M. May, *Infectious Diseases in Humans*: Oxford University Press, 1992.
- [2] R. Axtell, R. Axelrod, J. M. Epstein, and M. D. Cohon, "Aligning simulation models: a case study and results," *Computational and Mathematical Organization Theory*, vol. 1, 1996.
- [3] N. J. T. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications*, 2nd ed. New York: Oxford University Press, 1975.
- [4] L. A. Barakat, H. L. Quentzel, J. A. Jernigan, D. L. Kirschke, K. Griffith, S. M. Spear, K. Kelley, D. Barden, D. Mayo, and e. al., "Fatal inhalational anthrax in a 94-year-old Connecticut woman," *Journal of American Medical Association*, vol. 287, pp. 863-868, 2002.
- [5] J. G. Bartlett, J. Thomas V. Inglesby, and L. Borio, "Management of anthrax," *Clinical Infectious Disease*, pp. 851-858, 2002.
- [6] A. F. K. Brachman, et al., "Industrial inhalation anthrax," *Bacteriological Reviews*, vol. 30, pp. 646-657, 1966.
- [7] S. Braithwaite, D. Fridsma, and M. S. Roberts, "The cost-effectiveness of strategies to reduce mortality from an intentional release of aerosolized anthrax spores," School of Medicine, University of Pittsburgh, Pittsburgh, PA 2003.
- [8] R. Brookmeyer and N. Blades, "The statistical analysis of truncated data: application to the Sverdlovsk anthrax outbreak," *Biostatistics*, vol. 2, pp. 233-247, 2001.
- [9] D. Buckeridge, private communication, 2003.
- [10] K. Carley, N. Altman, B. Kaminsky, D. Nave, A. Yahja, 2004. *BioWar: A City-Scale Multi-Agent Network Model of Weaponized Biological Attacks*, Technical Report (CMU-ISRI-04-101). Pittsburgh, PA: CASOS, Carnegie Mellon University, available at <http://reports-archive.adm.cs.cmu.edu/isri2004.html>.
- [11] L.-C. Chen, B. Kaminsky, T. Tummino, K. M. Carley, E. Casman, D. Fridsma, A. Yahja, 2004 "Aligning Simulation Models of Smallpox Outbreaks," CASOS working paper, Carnegie Mellon University.
- [12] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. New York: John Wiley & Sons, 2000.
- [13] K. S. Griffith, P. Mead, G. L. Armstrong, J. Painter, K. A. Kelley, A. R. Hoffmaster, D. Mayo, D. Barden, R. Ridzon, U. Parasha, E. H. Teshale, J. Williams, S.

Noviello, J. F. Perz, E. E. Mast, D. L. Swerdlow, and J. L. Hadler, "Bioterrorism-related inhalational anthrax in an elderly woman, Connecticut, 2001," *Emerging Infectious Disease*, vol. 9, 2003.

[14] J. Guillemin, *Anthrax: The Investigation of a Deadly Outbreak*. Berkeley and Los Angeles, CA: University of California Press, 1999.

[15] T. V. Inglesby, T. O'Toole, D. A. Henderson, J. G. Barlett, M. S. Ascher, E. Eitzen, A. M. Friedlander, J. Gerberding, J. Hauer, J. McDade, M. T. Osterholm, G. Parker, T. M. Perl, P. K. Russell, and K. Tonat, "Anthrax as a biological weapon, 2002: updated recommendations for management," *Journal of American Medical Association*, vol. 287, pp. 2236-2252, 2002.

[16] J. A. Jernigan, D. S. Stephens, D. A. Ashford, C. Omenaca, M. S. Topiel, M. Galbraith, M. Tapper, T. L. Fisk, and e. al., "Bioterrorism-related inhalational anthrax: the first 10 cases reported in the United States," *Emerging Infectious Disease*, vol. 7, pp. 933-944, 2001.

[17] D. B. Jernigan, P. L. Raghunathan, B. P. Bell, R. Brechner, E. A. Bresnitz, and e. al., "Investigation of bioterrorism-related anthrax. United States, 2001: Epidemiologic findings," *Emerging Infectious Disease*, vol. 8, pp. 1019-1028, 2002.

[18] M. A. Louie, K. M. Carley, L. Haghshenass, J. C. Kunz, and R. E. Levitt, "Model Comparisons: Docking ORGAHEAD and SimVision," presented at Proceedings of NAACSOS conference, Pittsburgh, PA, 2003.

[19] A. B. Lawson, "Statistical Methods in Spatial Epidemiology", John Wiley & Sons, 2001.

[20] M. Meselson, J. Guillemin, M. Hugh-Jones, A. Langmuir, I. Popova, A. Shelokov, and O. Yampolskaya, "The Sverdlovsk anthrax outbreak of 1979," *Science*, vol. 266, pp. 1202-1208, 1994.

[21] W. C. Patrick, "Biological Terrorism and Aerosol Dissemination," *Politics and the Life Sciences*, vol. 15, pp. 208-210, 1996.

[22] L. M. Wein, D. L. Craft, and E. H. Kaplan, "Emergency response to an anthrax attack," *Proc. National Acad. of Sciences*, vol. 100, pp. 4346-4351, 2003.

[23] A. Yahja, K. M. Carley, D. Fridsma, E. Casman, N. Altman, B. Kaminsky, and D. Nave, 2003. "BioWar: scalable agent-based model of bioattacks," NAACSOS conference proceedings, Pittsburgh, 2003.

[24] A. Yahja and K. M. Carley, "WIZER: What-If Analyzer for Automated Social Model Space Exploration and Validation", NAACSOS Conference 2003, Pittsburgh, PA

TABLES

Population in the defined state	IPF	BioWar
Exposed	An input parameter based on the number of people taking prophylaxis in the US data and the number of people vaccinated in the Sverdlovsk data	The estimation is based on assumptions on wind speed, release height, release location and release mass and simulated data on geographic distribution of the population.
Incubation	An input parameter calculated based on the infection rate from the two empirical cases	Estimation is probabilistically based on agent's age and number of spores inhaled. The lognormal distribution randomly generates the duration of incubation period for each infected agent.
Prodromal Fulminant	Calculated based on the state transition probability calibrated from population level data of disease progression observed.	The lognormal distribution randomly generates the duration of prodromal and fulminant stages for each infected agent.
Death	Calculated based on the state transition probability calibrated from the number of deaths in the two empirical cases	An internal death probability of an individual agent determines if the agent will die or recover.
Recovery	Calculated based on the state transition probability calibrated from the number of recovery in the two empirical cases	

Table 1: Alignment of model components between BioWar and IPF based on stages of disease progression

Disease Stage	Mean , days	Standard Deviation , days
Incubation	2.4	0.71
Prodromal	0.85	0.35
Fulminant	0.34	0.35

Table 2a: The mean and standard deviation of the lognormal distribution for the three stages of anthrax (dose-independent case)

Disease Stage	Low Dose Mean, days	Medium Dose Mean, days	High Dose Mean, days
Incubation	2.7	2.4	1.4
Prodromal	0.99	0.85	0.61
Fulminant	0.41	0.34	0.16

Table 2b: The mean of the lognormal distribution for the three stages of anthrax (dose-dependent case). Low dose case corresponds for the less than 4000 spores inhaled, high dose – greater than 12000 spores inhaled, and medium case – between 4000 and 12000 spores inhaled.

Model parameters	Value
Simulation duration	400 days
Population of the city	148,000
Release mass	3000g (150g effective)
Dose dependency of the disease stage duration	Dose independent
Efficiency	0.05
Height of release	5 m (explosive release)
Release location	Municipal stadium (roughly 12,820 people are gathering inside the stadium)
Time of release	4pm (stadium full capacity)
Wind speed	4.617 m/sec
Treatment assumptions	People have a low initial probability being correctly diagnosed if they go to doctors since the early symptoms are similar to flu.
Spore resuspension and activity assumptions	Spores are not resuspended once they settle to the ground. Spores are only infective while suspended in air.

Table 3: Model parameters and assumptions for the BioWar scenarios

Data set	Data type	Exposed population	Infected population ²	Empirical infection rate ³	Simulated infection rate ⁴	Death rate ⁵	Stabilization time ⁶ (days)
US	Empirical	Unknown ⁷ .	11	0.10%	N.A.	45%	44
	IPF	Unknown	11	0.10%	N.A.	41%	32
	BioWar	28,757	2740	N.A.	9.5%	42%	42
Sverdlovsk	Empirical	Unknown ⁸	77	0.16%	N.A.	86%	66
	IPF	Unknown	77	0.16%	N.A.	86%	47
	BioWar	28,701	2779	N.A.	10%	86%	39

Table 4: A comparison of the results between BioWar and IPF with the empirical data sets

² The discrepancy in infected population between IPF and BioWar is due to the difference in the release mass of anthrax spores. IPF calibrates the infected population to empirical data and BioWar calculates it based on an attack scenario that the effective release mass is about 150 times of the Sverdlovsk case.

³ Empirical infection rate = infected population / the number of people taking anti-microbial prophylaxis or vaccinated.

⁴ Simulated infection rate = infected population / the exposed population. The exposed population refers to persons who are inhaled at least one anthrax spore.

⁵ Death rate = total number of deaths / infected population.

⁶ Stabilization time is the number of days that have elapsed when 99% of infected people either die or recover.

⁷ Approximately 10,300 persons completed a 60-day course of anti-microbial prophylaxis. This program is only applied to the people who met the following three factors: 1) the presence of an inhalational anthrax at a facility, 2) environmental specimens positive for *B. anthracis* in facilities along the path of a contaminated letter where aerosolization might have occurred, and 3) exposure to an air space known to be contaminated with aerosolized *B. anthracis* from an opened letter [5].

⁸ A voluntary immunization program vaccinated approximately 47,200 persons at least once. The voluntary immunization program using a live nonencapsulated spore vaccine was carried out for healthy persons 18 to 55 years old. Approximately 59,000 persons are eligible for the program and 80% were vaccinated at least once [14].

Figures

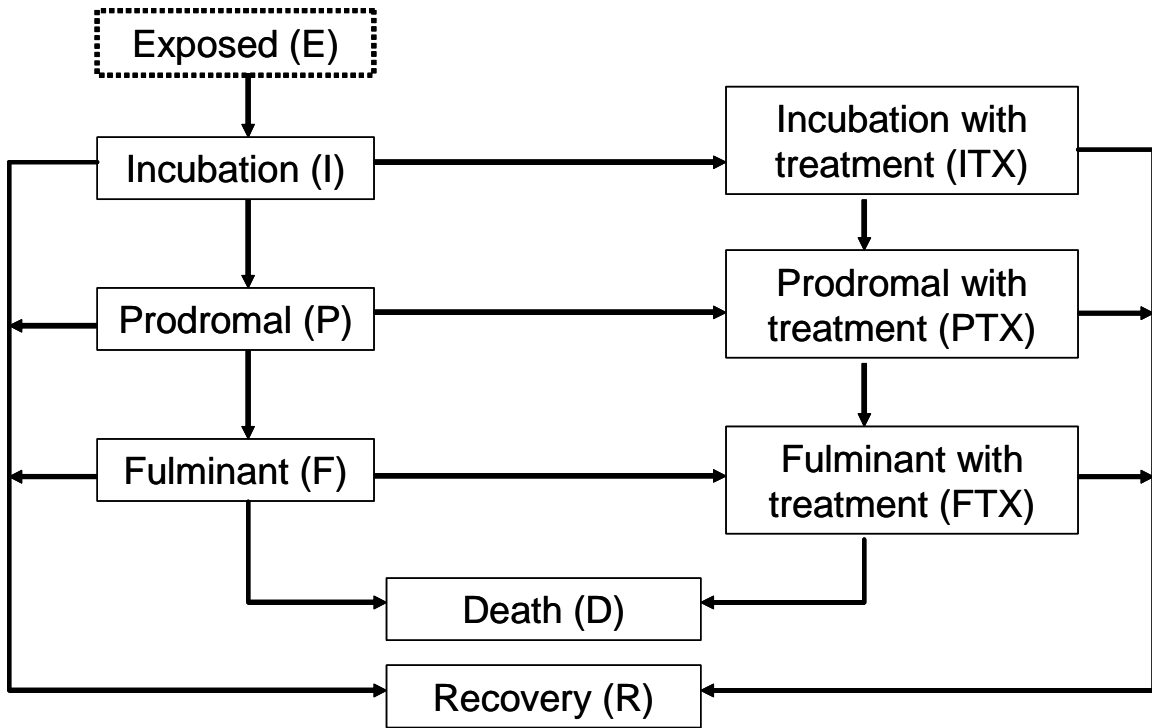


Figure 1: The IPF model

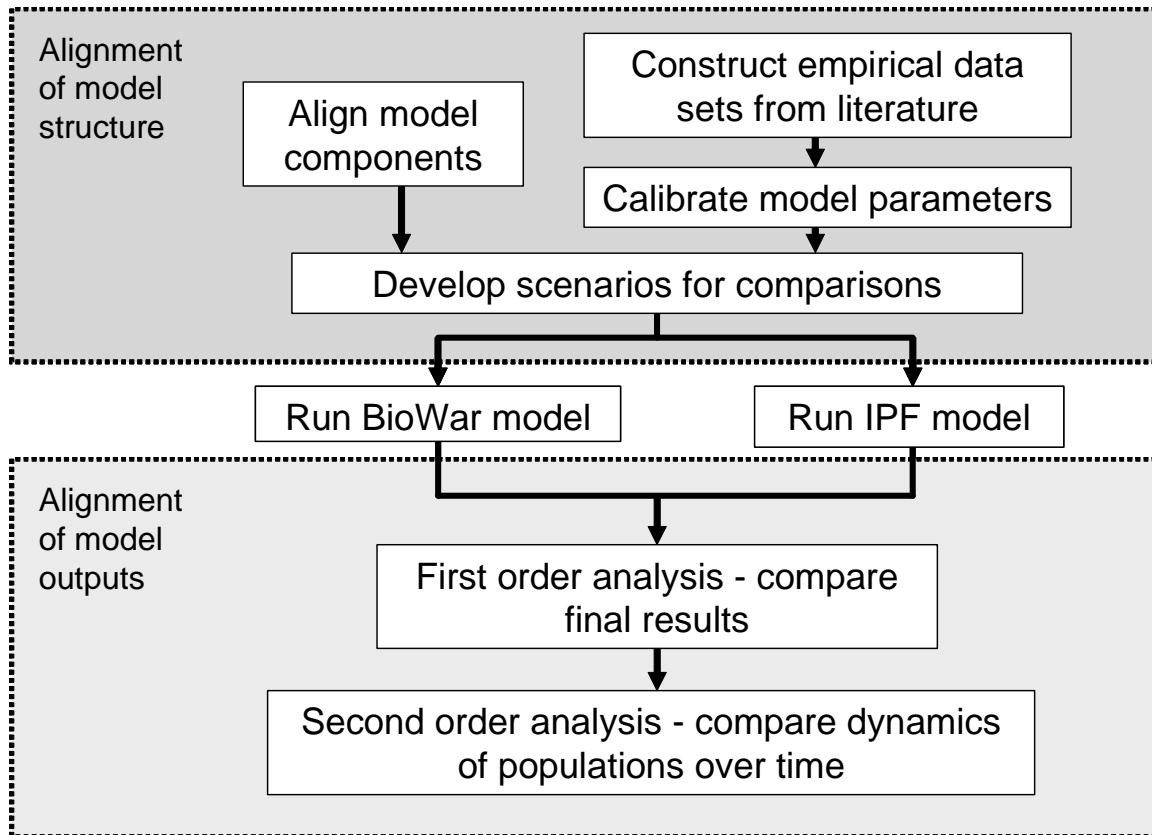


Figure 2: The process of model alignment

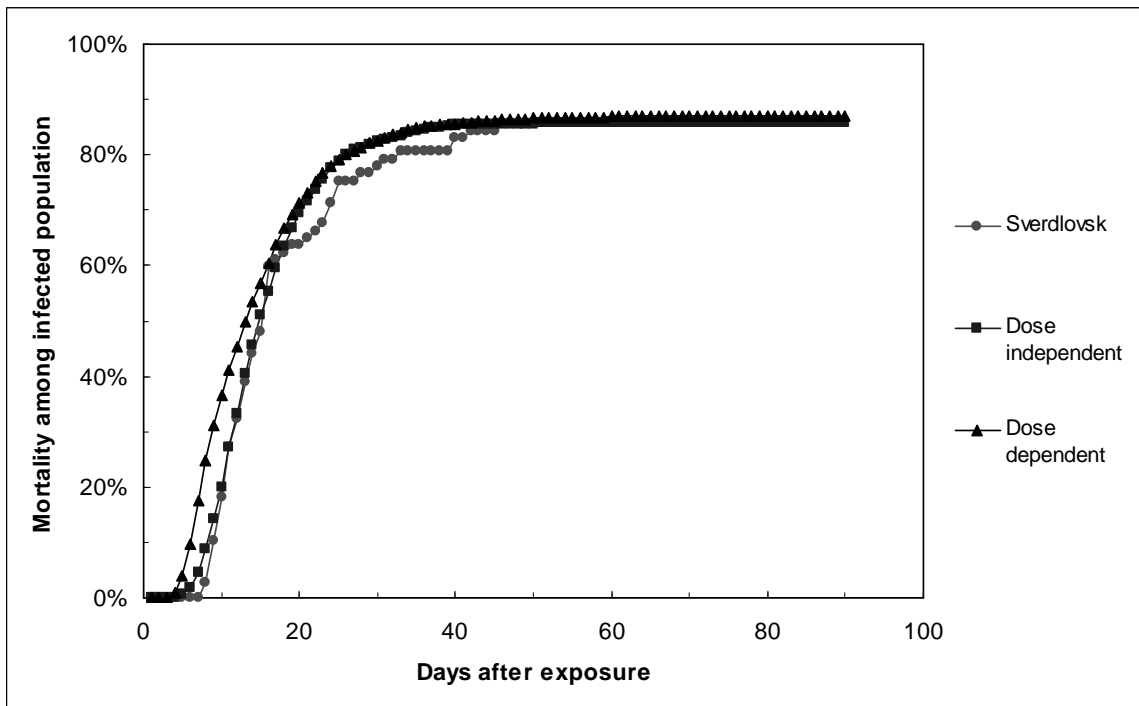


Figure 3: The comparison between BioWar scenarios with different dose dependency assumptions and the Sverdlovsk data

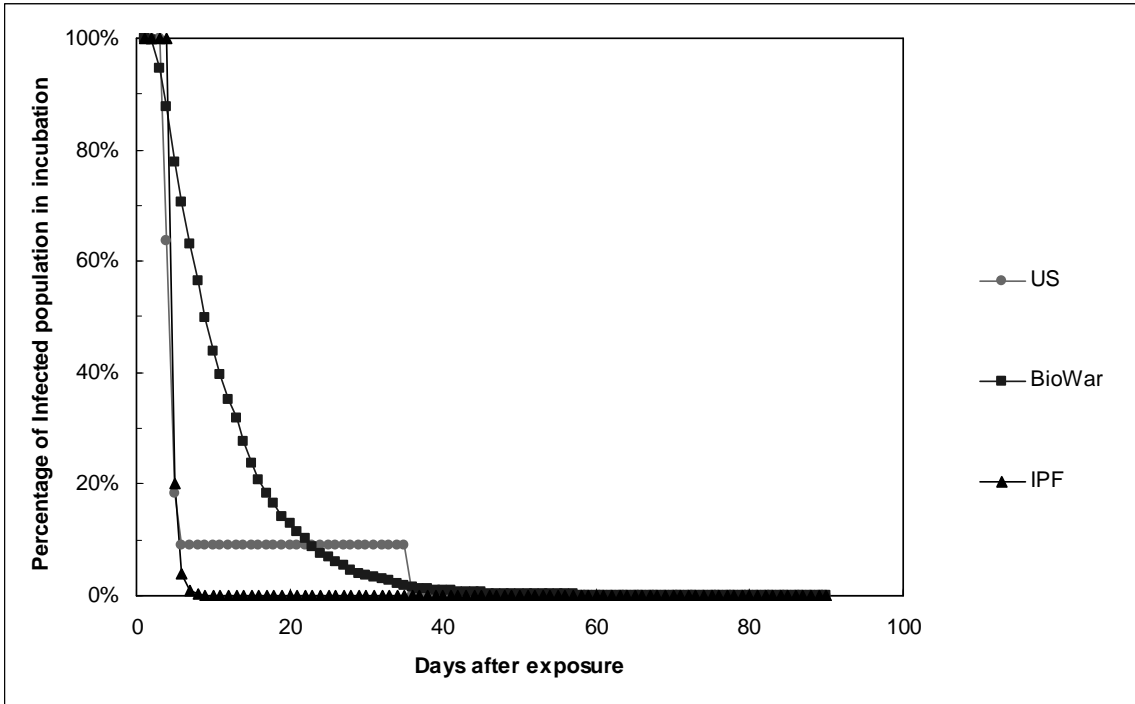


Figure 4: Comparison between IPF, BioWar and the US data for the percentage of infected population in the incubation stage

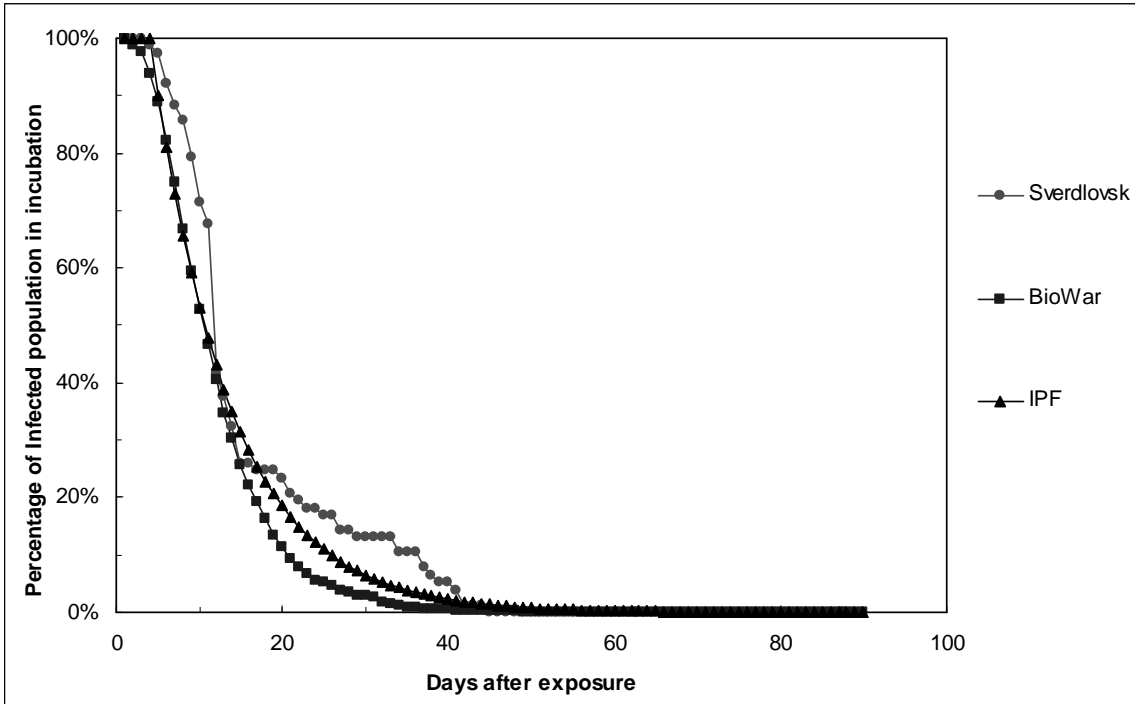


Figure 5: Comparison between IPF, BioWar and the Sverdlovsk data for the percentage of infected population in the incubation stage

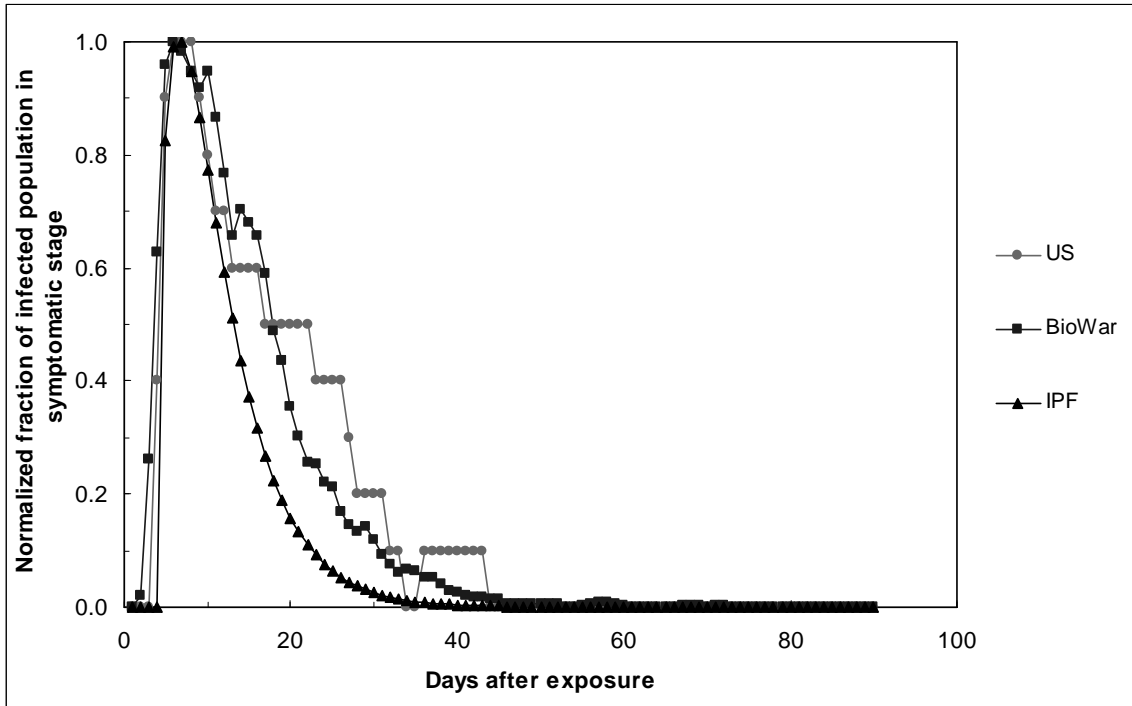


Figure 6: Comparison between IPF, BioWar and the US data for the normalized fraction of infected population in the symptomatic stage

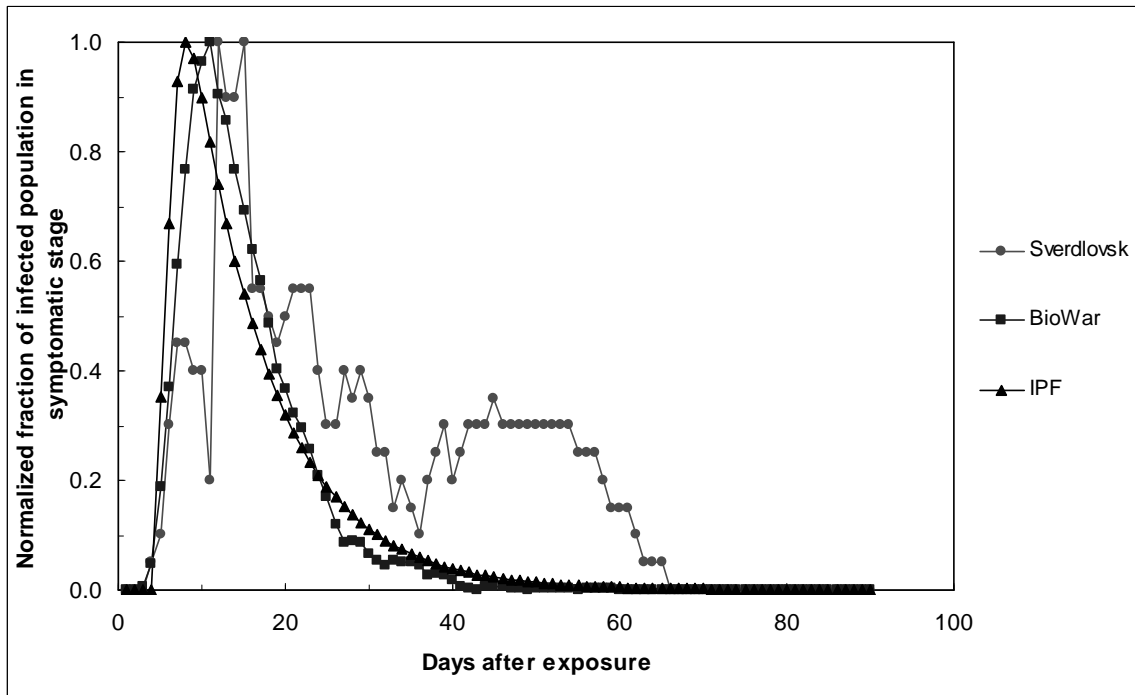


Figure 7: Comparison between IPF, BioWar and the Sverdlovsk data for the normalized fraction of infected population in the symptomatic stage

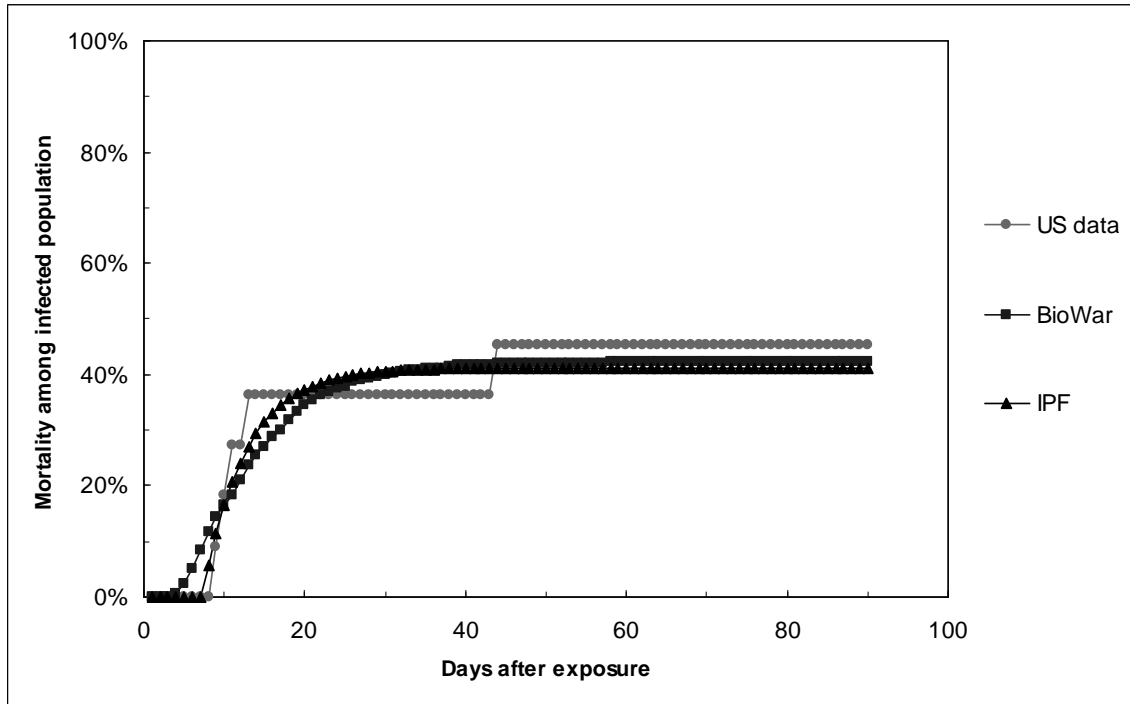


Figure 8: Mortality comparison between IPF, BioWar and the US data

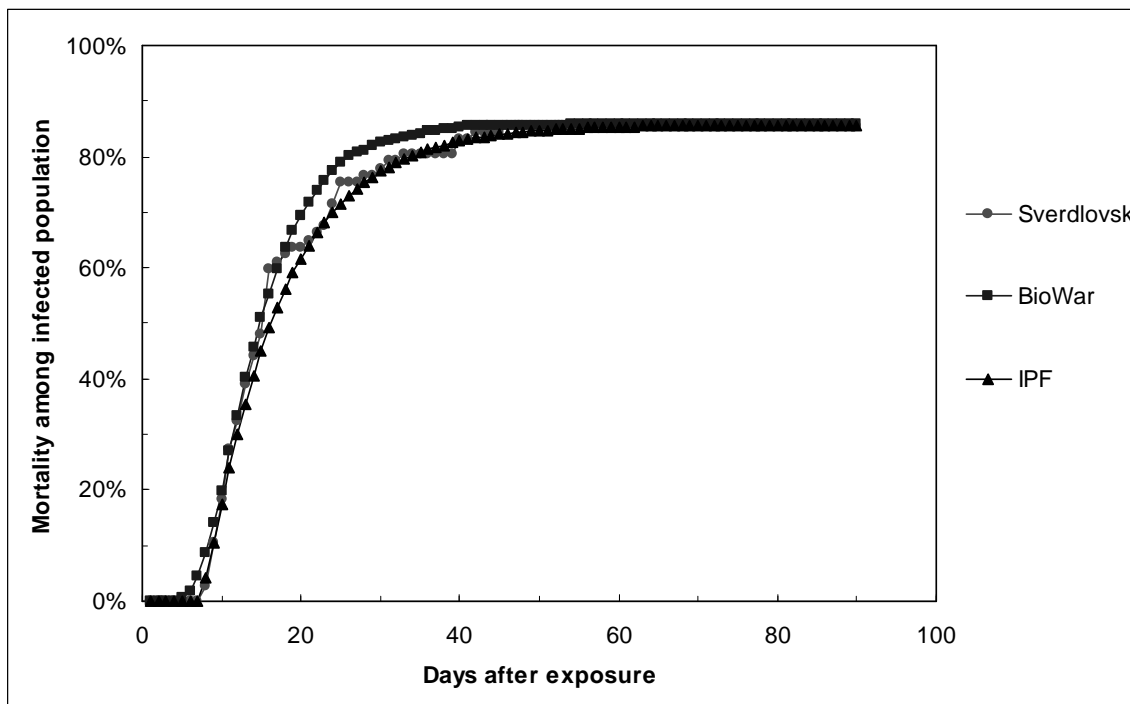


Figure 9: Mortality comparison between IPF, BioWar and the Sverdlovsk data

Appendix A: The IPF model

The total population exposed to anthrax spores N is divided into nine states: exposed but not yet infected (E), incubation (I), prodromal (P), fulminant (F), incubation with treatment (ITX), prodromal with treatment (PTX), fulminant with treatment (FTX), population that die (D), and population that recover (R). Each state is represented as a rectangular box in Figure 1.

$$N = E + I + P + F + ITX + PTX + FTX.$$

Infection rate, α , represents the fraction of exposed population infected after an attack. Transition probabilities are denoted as γ with two subscripts: the previous state and the current state. The changes of populations over time are described by equations (1).

$$\begin{aligned}
 I_0 &= \alpha E \\
 \frac{dI}{dt} &= -(\gamma_{I_ITX} + \gamma_{I_R} + \gamma_{I_P})I \\
 \frac{dITX}{dt} &= \gamma_{I_ITX}I - (\gamma_{ITX_PTX} + \gamma_{ITX_R})ITX \\
 \frac{dP}{dt} &= \gamma_{I_P}I - (\gamma_{P_PTX} + \gamma_{P_F} + \gamma_{P_R})P \\
 \frac{dPTX}{dt} &= \gamma_{P_PTX}P + \gamma_{ITX_PTX}ITX - (\gamma_{PTX_FTX} + \gamma_{PTX_R})PTX \\
 \frac{dF}{dt} &= \gamma_{P_F}P - (\gamma_{F_FTX} + \gamma_{F_D} + \gamma_{F_R})F \\
 \frac{dFTX}{dt} &= \gamma_{F_FTX}F + \gamma_{PTX_FTX}PTX - (\gamma_{FTX_D} + \gamma_{FTX_R})FTX \\
 \frac{dD}{dt} &= \gamma_{F_D}F + \gamma_{FTX_D}FTX \\
 \frac{dR}{dt} &= \gamma_{I_R}I + \gamma_{ITX_R}ITX + \gamma_{P_R}P + \gamma_{PTX_R}PTX + \gamma_{F_R}F + \gamma_{FTX_R}FTX
 \end{aligned} \tag{1}$$

APPENDIX B: CONSTRUCTION OF THE EMPIRICAL DATA SETS

The US data is based on the 2001 anthrax letter incidents in the United States [4-5, 13, 15-17]. There were eleven confirmed inhalational anthrax cases of whom five died. We collected the data set from existing literature to calculate the populations in the four stages of the disease progression: incubation, symptomatic, death, and recovery. Four cases in the US data have unknown incubation dates and we estimated the number in median days of incubation from available cases. The median of the incubation stage observed for the US mail attacks was four days, which is about 6-7 days shorter than that for the Sverdlovsk release. The date of incubation for the case of the 94-year-old Connecticut woman is estimated as the maximum possible number of days of incubation since the exact exposure date is unknown [4, 13].

The source of the Sverdlovsk data is based on published anthrax studies [14, 20]. The Sverdlovsk data has 77 confirmed inhalational anthrax cases and 66 deaths. We estimated the unknown data of disease stages in [20] based on their distributional estimates [8]. The actual number of days for recovery for individuals is not available in [20] but it was reported approximately 3 weeks hospital stay for survivors.