Extending an Existing Multifidelity Method to Simulate Rare Events in Systems of Multiple Tuberculosis Granulomas

David R. Wagner, Elsje Pienaar
Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

ABSTRACT: The higher mortality rate and more costly treatment of drug resistant tuberculosis than drug susceptible tuberculosis makes the emergence of drug resistant tuberculosis a serious threat to society. Simulating mutations leading to such a disease could aid in decision making regarding treatment. However, simulating such rare events is computationally expensive. Here, a previously developed multifidelity method applying Markov Chains and Monte Carlo Simulations to minimize this cost is extended to describe medically relevant systems containing multiple tuberculosis granulomas. Distributions containing probabilities for each possible number of mutant bacteria from randomly selected granulomas were created using the original multifidelity approach and combined in a pairwise manner until a system containing 40 granulomas was yielded. The resulting probability of systemwide drug resistance and frequencies for each possible total number of drug resistant mutants were then plotted. As the number of granulomas in the system increased, it was found that the overall probability of mutant bacteria also increased. In addition, it was observed that the frequency of occurrence for all possible numbers of resistant bacteria except zero increased with the number of granulomas. The frequency of zero resistant bacteria occurring dropped as the number of granulomas increased (reflecting observations in the probability plots). These results indicate that individuals with more granulomas have a higher risk of developing drug resistant tuberculosis and may assist in making decisions regarding treatment options.

KEYWORDS: Tuberculosis, multifidelity, rare events, drug resistance, granulomas

Introduction

Each year, there are approximately 10 million cases of tuberculosis worldwide [1] with 1.5 million associated deaths in 2014 alone [2]. Multidrug resistant tuberculosis (MDR-TB) is even more problematic due to its higher mortality rate [2] and the more intensive drug regimens it requires for treatment [1], [3]. Individuals infected with MDR-TB are resistant to isoniazid and rifampicin, two of the critical, first-line drugs, and are very difficult to cure unless resistance is identified early in the infection [4]. Unfortunately, it is difficult to study the occurrence of such drug resistant mutations due to the extremely low frequencies in which they occur. As the frequency of an event decreases, the number of observations required to reliably observe it increases. Thus, it is often time consuming and expensive to observe rare biological events experimentally. Computational models simulating such rare events are often used to reduce the amount of resources required to make useful observations [5]. However, as the complexity of the system being simulated increases, so too does the associated computational cost [6]. To combat this, certain methods have been used to optimize the allotment of computational resources [5], [6]. Here a previously developed approach optimized to model the occurrence of drug resistance in single tuberculosis granulomas will be expanded to simulate the same behavior in more medically relevant systems consisting of multiple granulomas.

Physical studies have been conducted in the past to gain knowledge regarding the behavior of tuberculosis granulomas. However, unique host characteristics have led to highly variable responses across different animal models as well as between different individuals belonging to the same model. Such varied outcomes have simply served to highlight the dynamic nature of tuberculosis granulomas and the delicate (but important) role of immune responses (such as the regulation and recruitment of neutrophils) [7], [8].

Tuberculosis granulomas have also been simulated using complex and highly detailed agent-based models such as *GranSim* [9]. *GranSim* works by discretizing a small section of lung parenchyma tissue (approximately equivalent to a 16 square millimeter section of tissue) and modeling macrophages, bacteria, and T cells as discrete "agents". Chemokines, cytokines, and, if desired, antibiotics are also considered and modeled as continuous throughout the discretized region of parenchyma. Interactions between these components (mimicking the natural dynamics found in tuberculosis granulomas) are then accounted for statistically to make predictions throughout successive time steps of the simulation [9]. Unfortunately, the complexity and

highly detailed nature of *GranSim* make it too computationally intensive to study rare events, such as the occurrence of drug resistance-conferring mutations, where many replicates are required for their statistical analysis [5].

As mentioned before, certain methods have been developed and implemented to minimize the computational cost of simulating rare events [6]. The model used as a basis here employs a "multifidelity" approach where a bacterial population trajectory within a single granuloma is first calculated using the high-fidelity (meaning highly accurate) *GranSim* model. This bacterial population trajectory from *GranSim* is then fed into medium and low-fidelity models based on Monte Carlo Simulations and Markov Chains respectively. Both models step through each bacteria division and death recorded in the high-fidelity population trajectory. At each step, probabilities associated with mutations spontaneously occurring as well as mutant bacteria dividing or dying are recalculated and applied to determine the overall probabilities associated with every possible number of mutant bacteria within the granuloma at the current time point. The advantage of this method is that only a single high-fidelity output is needed by the much less computationally intensive medium and low-fidelity models to calculate probability distributions for the occurrence of drug resistance (rather than many replicates of the high-fidelity model) [5].

To fully define the scope of this project, one interesting feature of tuberculosis infections must be clarified. The drug resistance discussed here refers to resistance conferred strictly by genetic mutations, not by epigenetic or metabolic states. That is, only methods of drug resistance caused by alterations in the bacterium's genetic sequence, such as those discussed by Palomino [10], are considered. Other methods of reducing drug efficacy not caused by genetic mutations (e.g., metabolic state [7], granuloma heterogeneity [8], or lack of drug diffusivity [11]) are not considered in this study.

Adapting the existing model described above to scenarios with more than one tuberculosis granuloma will improve the accuracy of results regarding the occurrence of drug resistance as individuals infected with tuberculosis generally develop more than one granuloma [12]. Applying the multi-fidelity approach to systems of multiple granulomas may also aid in decision-making regarding tuberculosis treatment options since the probability of pre-existing drug resistance could be estimated as a function of the time since the infection started [4], [5]. In general, extending the initial multi-fidelity model to more accurate cases with multiple granulomas will provide more insight into the occurrence of drug resistance in tuberculosis infections.

Methods

To model systems containing multiple tuberculosis granulomas, an algorithm was developed to combine the probability distributions outputted by the previously discussed multifidelity model. These outputs (whether for the low-fidelity or medium-fidelity model) take the forms of matrices where each element represents the probability of a certain number of resistant bacteria at a certain time (specified by its location within the matrix, see Figure 1).

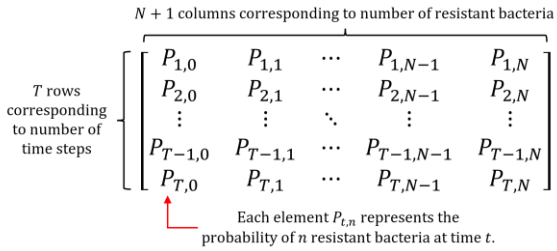

Figure 1. Illustration of probability distribution outputs from both medium and low fidelity models. There are as many rows as there are time steps and as many columns as the maximum number of drug resistant bacteria plus one (to account for the probability of *zero* resistant bacteria). Thus, the row index, t, corresponds to the time while the column index, n, corresponds to the number of resistant bacteria.

To combine the probability distributions for multiple granulomas, probabilities for every possible combination of all viable numbers of resistant bacteria must be individually combined. That is, the individual probabilities for every possible way in which n drug resistant bacteria can be separated among m granulomas must be summed. The result of this summation is the overall probability of n drug resistant bacteria within the system of multiple granulomas. Of course, this procedure must be repeated for all possible numbers of resistant bacteria within the system (which would range from zero to the total population of bacteria within the system) and for all time steps. An example algorithm illustrating this method applied to a system of two granulomas is presented in Figure 2.

```
Iterate through all time points T
For t from 0 to T:
    Iterate through all possible numbers of resistant bacteria N
    For n from 0 to N:
        Reset probability counter
        Prob = 0
        Iterate through all distribution of n resistant bacteria between 2 granulomas
        For i from 0 to n:
            Add the next probability to the counter
            Prob = Prob + ProbA(n − i) × ProbB(i)
        End
    End
End
```

Figure 2. Algorithm demonstrating the combination of two probability distributions (representing two granulomas). Every possible distribution of all present mutant bacteria is explored for all time points t from 1 to T. Here, N represents the total number of bacteria within the two-granuloma system and n iterates successively through each possible number of mutant bacteria within the system to calculate its associated probability (taking on values from 0 to N). For each value n takes on, the probabilities for all possible distributions of n mutant bacteria between the two granulomas are summed into the "Prob" variable ("ProbA" and "ProbB" represent the probabilities their associated number of mutant bacteria within granulomas A and B respectively).

One issue can arise using this algorithm, however. Since the general form of the algorithm for two granulomas (depicted in Figure 2) begins with all resistant bacteria in the first granuloma and successively "shifts" them to the second granuloma some of the distributions of resistant bacteria will not be viable. For example, if each of the two granulomas has 100 bacteria, the Figure 2 algorithm will begin by allotting all 200 possible mutants into the first granuloma and then progressively pass them to the second granuloma. Initially, this problem was solved by concatenating both input matrices with columns containing zeros so any probability calls that would be out of bounds of the original matrix will simply return a zero (indicating a probability of zero and hence not altering the mathematics).

Although the original algorithm worked very well, its efficiency with respect to both time and computer memory was unsatisfactory. To improve the run times and memory usage, a new algorithm was developed that altered the bounds of the iterations through each possible distribution of resistant bacteria so that unviable combinations are not considered rather than simply assigned probabilities of zero (see Figure 3). Such combinations are avoided by recognizing how they occur as the number of possible resistant bacteria, n, is iterated from 0 to N (the total population of the system). As n increases from 0 to the maximum population of the smaller granuloma, no unviable combinations of n between the two granulomas can be formulated. Therefore, for each n in this range, no restrictions are necessary when formulating possible combinations (see region 1 of Figure 3). As an example, if a system contained two granulomas where one has a population of 100 bacteria and the other 200, this range would span from 0 to 100 bacteria. Once n exceeds the population of the smaller granuloma, unviable combinations will occur. However, depending on the sizes of the granulomas, unviable combinations with respect to both granulomas may or may not yet be present. In the general case where the two granulomas may have different populations there is a brief range between the maximum population of the smaller granuloma and the maximum population of the larger granuloma where unviable combinations of n resistant bacteria are only unviable with respect to the smaller of the two granulomas. To avoid such combinations in this range, all n resistant bacteria can be initially placed within the larger granuloma and the successive "shifting" of said resistant bacteria into the smaller granuloma can be prematurely halted when its maximum population has been attained (see region 2 of Figure 3). Revisiting the above illustration, this region would account for a range between 101 and 200 bacteria. As n exceeds the population of the larger granuloma, unviable combinations are certain to occur with respect to both granulomas. To avoid all unviable combinations here, the iterative "shifting" of n resistant bacteria between the two granulomas can be initiated with the larger granuloma "full" and all remaining resistant bacteria (equal to n minus the population of the larger granuloma) within the smaller granuloma. As with the previous range, this iterative process will halt as soon as the maximum population of the smaller granuloma is reached (see region 3 of Figure 3). Using the above example, this region corresponds to a range from 201 to 300 bacteria.

```
Iterate through all time points T
For t from 0 to T:
  Region 1
    Iterate through all possible numbers of resistant bacteria that do not exceed either population
    For n from 0 to maxPopA:
        Prob = 0
        Iterate through all distribution of n resistant bacteria between 2 granulomas
        For i from 0 to n:
            Prob = Prob + ProbA(n − i) × ProbB(i)
        End
    End
  Region 2
    Iterate through all possible numbers of resistant bacteria that exceed the population of A but not B
    For n from maxPopA +1 to maxPopB:
        Prob = 0
        Iterate through all distribution of n resistant bacteria between 2 granulomas
        For i from 0 to maxPopA:
            Prob = Prob + ProbA(n − i) × ProbB(i)
        End
    End
  Region 3
    Iterate through all possible numbers of resistant bacteria that exceed the populations of A and B
    For n from maxPopB + 1 to N:
        Prob = 0
        Iterate through all distribution of n resistant bacteria between 2 granulomas
        For i from n − maxPopB to maxPopA:
            Prob = Prob + ProbA(n − i) × ProbB(i)
        End
    End
End
```

Figure 3. More efficient algorithm combining probability distributions representing two granulomas A and B. The terms "maxPopA" and "maxPopB" represent the maximum populations within granulomas A and B respectively. Here A is *always* the smaller of the two granulomas (it has a smaller maximum population than that of granuloma B).

The algorithms discussed above were developed with respect to systems containing only two granulomas. Perhaps unsurprisingly, the simultaneous combination of multiple probability distributions becomes exponentially more convoluted as the number of distributions increases. Therefore, for the sake of simplicity, each probability distribution was combined in a repetitive, pairwise manner until the desired number of distributions was merged (see Figure 4).

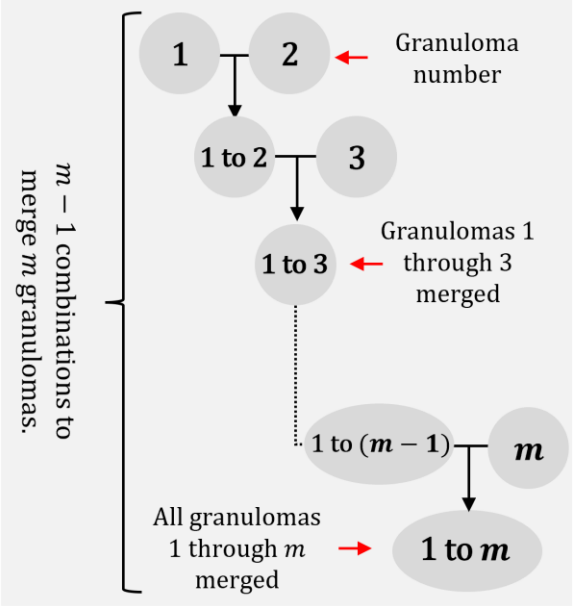

Figure 4. Pairwise combination of granuloma probability distributions. To avoid the complexity associated with combining the distributions of more than two granulomas simultaneously, each distribution was combined in a pairwise manner. That is, in a system of m granulomas, each distribution is successively "collapsed" across $m - 1$ steps into one single aggregate distribution.

With the above algorithms developed, simulations containing up to 40 granulomas were performed. A maximum of 40 granulomas was chosen to approximate the experimentally observed median number of granulomas in cynomolgus macaques [14]. Each of the 40 simulations was performed separately using the original multifidelity model and randomly drawn population data from a bank of 114 high-fidelity model outputs. Probability distributions resulting from all simulations were then combined using the method described previously. A mutation frequency of 2.56×10^{-8} was used to reflect the estimated observed drug-resistance mutation rates in tuberculosis responding to treatment by isoniazid [13]. To calculate the probability of resistance for the system of granulomas, the overall probability of zero resistant bacteria for all time points was subtracted from one. The result of this operation yields the probability of not having zero resistant bacteria (e.g., the probability that *any* resistance is present in the system). Frequencies corresponding to each possible number of resistant bacteria were also plotted at the time point where the first randomly selected granuloma reached its maximum population.

Results

Probability of Resistance over Time

As described above, the algorithm was applied to model a system containing 40 separate granulomas over a 200-day period. To provide appropriate statistical power, 10 replicates of this simulation were performed. The behavior of the probability for drug resistance as a function of time is displayed in Figure 5A. Results from systems containing 1, 5, 10, 20, 30, and 40 granulomas are depicted to better illustrate how the behavior of drug resistance varies as the number of granulomas is increased. Each darkened line represents the averages across all 10 replicates while the corresponding shaded regions denote the associated 95% confidence intervals. To give context to the probability distributions, bacterial populations for the same systems of granulomas are also plotted along the 200-day interval (Figure 5B). As before, the dark lines correspond to the average populations of the 10 replicates and the shaded regions illustrate the 95% confidence intervals.

The behavior of the probability of drug resistance displayed in Figure 5A generally agrees with the expected results based on intuition. As the number of granulomas increases, so too does the probability that resistance will occur somewhere within the system. Such an explanation also accounts for the remarkable similarity between the profiles for the probability of resistance (Figure 5A) and the bacterial population (Figure 5B). As the population of the system rises, the number of opportunities for drug resistance to occur increases as well. Likewise, when the population falls (as it does in this case once the adaptive immune system engages [11]), many of the mutant bacteria will be killed off along with their wild type companions. This direct relationship between the probability of resistance and the system population appears to hold extremely well across most of the 200-day period. Towards the end of the 200 days, however, the probability of resistance within the systems begins to increase slowly while the population remains stagnant. The continual bacterial "turnover" may explain this observation. As the bacteria within the granulomas divide and die over time, more opportunities are provided for either the division of mutant bacteria already present or for the appearance of a mutant daughter cell after the division of a wild type parent cell.

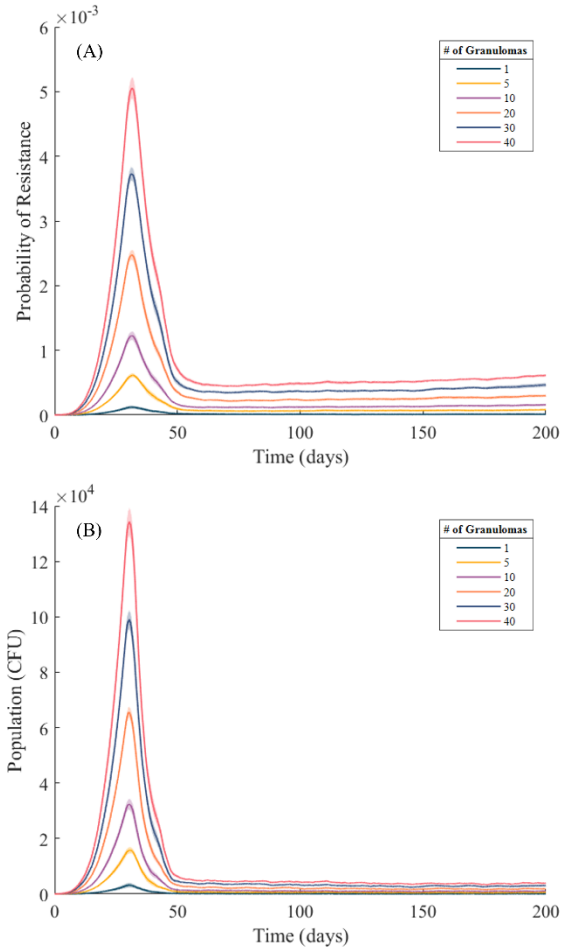

Figure 5. Probability distributions and population profiles for systems containing various numbers of tuberculosis granulomas. (A) Probability distributions for drug resistance as a function of time across a 200-day period. Probabilities are shown for multiple groups of granulomas (including up to 40) to explicitly display the differences in behavior as the number of granulomas increases. Darkened lines correspond to the average probability distributions across all 10 replicates and the shaded regions denote their 95% confidence intervals. (B) Systemwide bacterial populations as a function of time across a 200-day period. Populations for multiple systems containing up to 40 granulomas are shown to better illustrate the behavior as the number of granulomas is increased. Darkened lines denote the average populations across 10 replicates while the shaded regions illustrate the 95% confidence intervals.

Probability of Resistance vs. Number of Granulomas

To gain a better understanding of the precise nature of the relationship between the probability of drug resistance and the number of granulomas, the probability of resistance is plotted as a function of the number of granulomas in Figure 6. The figure displayed here mirrors Figures 5A and B above by considering systems containing the same numbers of granulomas (1, 5, 10, 20, 30, and 40). Separate plots are made across multiple time points representing 40, 80, and 200 days after the granulomas initially form to better evaluate this relationship. Linear regression models were fitted to each of the time datasets and the quality of these models was evaluated.

Although the rather evenly spaced probability of drug resistance curves displayed in Figure 5A hint to the existence of a linear relationship between the probability of resistance and the number of granulomas, the appearance of the data and corresponding models in Figure 6 confirms the linear nature of this relationship. For all the time points tested, the corresponding linear fits have coefficient of determination (R^2) values above 0.97, an indication that the linear models developed describe the variation in the data very well. It can also be noted that the behavior of the slopes of the linear fits agrees well with what has been observed in Figure 5A. The slope of the probability of resistance line at 40 days exceeds the slopes observed at both 80 and 200 days. Such behavior is explained by observing both the enormous "spike" in probability for drug resistance that occurs between 10 and 50 days and the relative "plateau" in probability existing afterwards. Greater separation between the probability curves for each system of granulomas in the spike region (relative to the plateau region) results in a greater slope of the line connecting the curves (shown in Figure 6). To better observe this, compare the separation between the curves in Figure 5A at both 40 and 80 days as well as the slopes of their corresponding lines in Figure 6. It is also interesting to note that the slight increase in the probability of drug resistance that occurs after about 150 days (see Figure 5A) is prominently displayed in Figure 6 as the slope of the linear fit corresponding to 200 days is noticeably higher than the slope for 80 days.

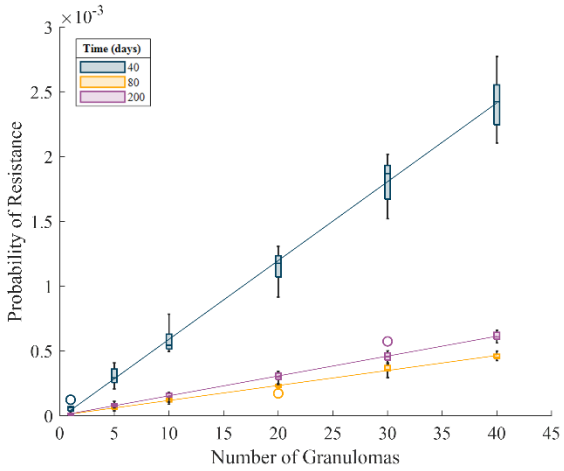

Figure 6. Probability of drug resistance displayed as a function of the number of granulomas at multiple time points (corresponding to 40, 80, and 200 days). The linear relationship between probability of resistance and number of granulomas is confirmed by the evaluation of the additional linear regression fits. All linear regression fits have coefficient of determination (R^2) values above 0.97, indicating an excellent description of the variation in the probabilities as a function of the number of granulomas.

Probabilities for Different Numbers of Mutant Bacteria

In the previous sections, consideration has only been given to the probability that drug resistance is present within the system of granulomas. That is, drug resistance has been treated as "binary" where it is either present or absent. Here, probabilities for various numbers of resistant bacteria within the system are considered. Probabilities are plotted as a function of the number of resistant bacteria for systems consisting of 1, 20, and 40 granulomas (Figure 7). All profiles are plotted at the time corresponding to the maximum population of the largest system of granulomas. In this case, the system of 40 granulomas experienced a maximum population at 30.4 days. Due to the significant amount of time required for this simulation (almost 22 hours), no replicates were performed, and the data displayed in Figure 7 were collected independently from the data used in producing Figures 5 and 6.

Looking at Figure 7, the probabilities for greater than zero mutant bacteria generally increase as the number of granulomas within the system increases. However, the probability of zero resistant bacteria occurring decreases as the number of granulomas is increased (illustrated in the expanded inset figure). Such results are rather intuitive as the presence of

more granulomas (and hence a higher overall system population) will naturally increase the probabilities associated with at least *some* resistant bacteria being present while decreasing the probability associated with *absolutely no* resistant bacteria being present.

Besides the relatively expected results regarding the differences between the probability profiles, other, more interesting behavior is observed as the number of granulomas is increased. In the curve representing the system of one granuloma, the probability suddenly drops at about 5000 mutant bacteria (its maximum population). The curves for 20 and 40 granulomas also experience a steep, "steplike" drop at around this time point. However, since their populations far exceed this range of bacteria, their curves continue after this drop. In fact, while it is not shown in Figure 7, it was determined that the curves for the systems containing 20 and 40 granulomas experience exactly 20 and 40 of these "steps" respectively. These steps are a direct result of the process of combining the granulomas. As the probability distributions for the granulomas are merged, probabilities in regions of significant overlap with respect to the population will be increased only slightly (look at the range from 0 to about 5000 bacteria in Figure 7). In the region spanning the maximum population of the larger granuloma to the maximum population of the total system, the probability profile will also increase. However, since probabilities this region represent only scenarios where a significant number of bacteria in both granulomas harbor resistance, these probabilities will be a few orders of magnitude lower than what was found in the previous region. This pattern of regional probabilities for resistance repeats as the number of granulomas is increased and leads to this "descending staircase" of probabilities which each "step" corresponding to the probabilities that all the component granulomas harbor drug resistance.

The presence of the smaller "steps" and "spikes" within the profiles displayed in Figure 7 can also be explained by the systematic combination of probability distributions. As was mentioned earlier, the upward "shift" in the probability profiles that occurs as the number of granulomas is increased can be explained by the "summation" of the overlapping of the probability profiles for each of the randomly selected granulomas. Since the possible granulomas are not identical in size, this region of overlap will not be universal. Rather, there will be small regions where only some of the probability distributions belonging to the individual granulomas overlap with each other. As a result, small discontinuities will appear at the number of mutant bacteria that correspond to the maximum populations of the component granulomas. The nature of this discontinuity, whether a "spike" or "step", is dependent on the behavior of the probability profile for the individual component granulomas. That is, some of the profiles end in spikes while others end in steps, resulting in a unique series of spikes and steps in the combined profile. Unlike the larger "steps" discussed in the previous paragraph, the number of these small discontinuities does not always correspond to the number of granulomas in the combined profile. This is likely a result of the random nature in which the granulomas are selected. If the same granuloma is selected multiple times, its unique discontinuity (whether a spike or step) will always appear at the same point corresponding to its maximum population. Thus, multiple discontinuities can appear as one and lead to a situation where it appears that there are less discontinuities than the number of granulomas.

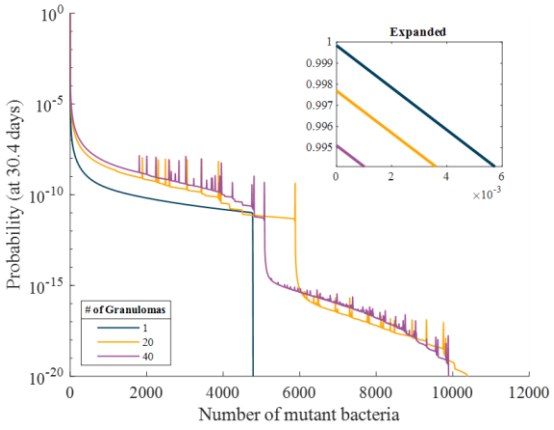

Figure 7. Probability of occurrence for different numbers of mutant bacteria. Plots for systems containing 1, 20, and 40 granulomas are shown at 30.4 days, the time corresponding to the maximum population of the 40-granuloma system. To better illustrate the detail where the plotted lines intersect the frequency axis, an expanded insert figure is displayed as well. Note that no replicates were performed due to the significant amount of time required for each simulation and that the data used to produce this figure was collected independently from the data used in Figures 5 and 6 above.

Discussion

As has been stated before, the goal of this research has been to adapt and apply a pre-existing multifidelity model of the emergence of drug resistance in tuberculosis granulomas to systems consisting of multiple granulomas. The development of the algorithms performing this task was described along with the relevant theory regarding the process of combining probability distributions. These algorithms were applied to systems consisting of up to 40 granulomas with a mutation frequency set to approximate the emergence of resistance to the antibiotic isoniazid (a common choice in the treatment of tuberculosis). Results were attained by plotting both the overall probability of resistance and the system populations as a function of time. Probability of resistance was also evaluated as a function of the number of granulomas by the fitting linear regression models. Lastly, probabilities for the presence of all possible numbers of drug resistant bacteria were explored graphically.

By plotting both probabilities of drug resistance and granuloma populations as functions of time, it was observed that the probability of resistance within systems of varying numbers of granulomas behave similarly to the populations of the corresponding granuloma systems. An exception to this general rule is the slight increase in drug resistance observed toward the end of the 200-day period caused by the continual "turnover" in the bacterial population. The linear relationship of the probability of resistance as a function of the number of granulomas was also demonstrated. However, it was observed the slope of this linear relationships changes with time. Thus, it can be concluded that the probability of resistance is a function of both time and the number of granulomas. Additionally, it was shown that probabilities for each possible number of resistant bacteria vary proportionally with the number of granulomas and inversely with respect to the number of resistant bacteria.

Although the goal driving the development of this research was to improve the applicability and relevancy of the multifidelity model, some limitations are still present. For example, while the mutant, drug resistant bacteria can be considered to have associated losses in relative fitness, the deterministic population trajectories remain unaffected. That is, the dynamics of the granuloma populations are unaffected by the differences in fitness between both wild type and mutant bacteria. The loss in fitness is only considered in the calculations regarding the probability of mutant bacterial birth. Such an assumption does not accurately reflect reality as the loss in fitness of mutant bacteria within a granuloma will alter the population trajectory of that granuloma with respect to that of an identical granuloma containing only wild type bacteria [5]. In addition to this, all granulomas are modeled as forming at the exact same time. No consideration is made for the process of bacterial dissemination. Since numerous waves of dissemination are common and lead to the development of new granulomas, assuming all granulomas form at the same time predisposes the results of the simulation to significant inaccuracies [8].

While the nature of the multifidelity model does not allow for the assumption regarding the effect of relative fitness on the granuloma population dynamics to be repealed, some modifications can be made in the future to improve the applicability of the model developed here. As described earlier, it is a rather poor assumption that all granulomas are initiated at the same time. Perhaps a method of combining granulomas in the future could be made to account for the process of granuloma seeding. Also, all the granulomas used here have been randomly pulled from a set of 114 unique *GranSim* outputs. Adding to this bank of high-fidelity outputs to feed into the multifidelity model may aid in providing a wider range of outcomes.

Acknowledgements

Significant amounts of advice and technical support were provided by Dr. Pienaar and her team of graduate students as well as Purdue's Summer Undergraduate Research Fellowship (SURF) team of graduate assistants. I thank them for their time and willingness to help. Funding for this research project was provided by the Purdue SURF program. However, this material is based upon work supported by the National Science Foundation under Grant No. 2143866 to E. Pienaar. I thank both Purdue SURF and the NSF for their financial support.

References

[1] World Health Organization. “Tuberculosis.” WHO.int. https://www.who.int/news-room/fact-sheets/detail/tuberculosis (accessed June 5, 2022).

[2] M. K. Cowan and H. Smith, “Infectious diseases affecting the respiratory system” in *Microbiology: A Systems Approach*, 5th ed. New York, NY, USA: McGraw Hill, 2018, ch. 21, sec. 5, pp. 630-634.

[3] Centers for Disease Control and Prevention. “Drug-Resistant TB.” CDC.gov. https://www.cdc.gov/tb/topic/drtb/default.htm (accessed June 5, 2022).

[4] K. J. Seung, S. Keshavjee, and M. L. Rich, “Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis,” *Cold Spring Harb. Perspect. Med.*, vol. 5, no. 9, Sept. 2015, doi: 10.1101/cshperspect.a017863.

[5] E. Pienaar, “Multifidelity analysis for predicting rare events in stochastic computational models of complex biological systems,” *Biomed. Eng. And Comput. Biol.*, vol. 9, no. 1, pp. 1-11, June 2018, doi: 10.1177/1179597218790253.

[6] R. M. Donovan, et al., “Unbiased rare event sampling in spatial stochastic systems biology models using a weighted ensemble of trajectories,” *PloS Comput. Biol.*, vol. 12, no. 2, Feb. 2016, doi: 10.1371/journal.pcbi.1004611.

[7] A. Lenaerts, C. E. Barry, and V. Dartois, “Heterogeneity in tuberculosis pathology, microenvironments and therapeutic responses,” *Immunol. Rev.*, vol. 264, no. 1, pp. 288-307, 2015, doi: 10.1111/imr.12252.

[8] A. M. Cadena, S. M. Fortune, and J. L. Flynn, “Heterogeneity in tuberculosis,” *Nat. Rev. Immunol.*, vol. 17, no. 11, pp. 691-702. Nov. 2017, doi: 10.1038/nri.2017.69.

[9] Kirschner Lab of Microbiology and Immunology at University of Michigan. “GranSim.” Umich.edu. http://malthus.micro.med.umich.edu/GranSim/ (accessed May 22, 2022).

[10] J. C. Palomino and A. Martin, “Drug resistance mechanisms in *Mycobacterium tuberculosis*,” *Antibiotics*, vol. 3, pp. 317-340. July 2014, doi: 10.3390/antibiotics3030317.

[11] E. Pienaar, et al., “A computational tool integrating host immunity with antibiotic dynamics to study tuberculosis treatment,” *J. Theor. Biol.*, vol. 376, pp. 166-179. Feb. 2015, doi: 10.1016/j.jtbi.2014.11.021.

[12] H. C. Warsinske, R. M. DiFazio, J. J. Linderman, J. L. Flynn, and D. E. Kirschner, “Identifying mechanisms driving formation of granuloma-associated fibrosis during *Mycobacterium tuberculosis* infection,” *J. Theor. Biol.*, vol. 429, pp. 1-17. Sept. 2017, doi: 10.1016/j.jtbi.2017.06.017.

[13] H. L. David, “Probability distribution of drug-resistant mutants in unselected populations of Mycobacterium tuberculosis,” *Appl. Microbiol.*, vol. 20, no. 5, pp. 810-814. Aug. 1970, doi: 10.1128/am.20.5.810-814.1970

[14] P. L. Lin, et al., “Radiologic responses in cynomolgus macaques for assessing tuberculosis chemotherapy regimens,” *Amer. Soc. for Micro.*, vol. 57, no. 9. Jun. 2013, doi: 10.1128/AAC.00277-13